



UNIVERSIDADE FEDERAL DO ESPÍRITO SANTO
CENTRO TECNOLÓGICO
DEPARTAMENTO DE ENGENHARIA ELÉTRICA
PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA ELÉTRICA

FLAVIO DA SILVA VITORINO GOMES

**MODELOS DE SÉRIES TEMPORAIS
PARA PREVISÃO DE NÍVEL DE
LÍQUIDOS EM CADINHO DE
ALTOS-FORNOS**

VITÓRIA-ES
2016

FLAVIO DA SILVA VITORINO GOMES

**MODELOS DE SÉRIES TEMPORAIS
PARA PREVISÃO DE NÍVEL DE
LÍQUIDOS EM CADINHO DE
ALTOS-FORNOS**

Tese de doutorado apresentada ao Programa de Pós-Graduação em Engenharia Elétrica da Universidade Federal do Espírito Santo como parte dos requisitos para a obtenção do título de Doutor em Engenharia Elétrica.

Área de concentração: Sistemas de Controle

ORIENTADOR: Prof. Dr. José Leandro Félix Salles

VITÓRIA-ES
2016

FLAVIO DA SILVA VITORINO GOMES

**MODELOS DE SÉRIES TEMPORAIS
PARA PREVISÃO DE NÍVEL DE
LÍQUIDOS EM CADINHO DE
ALTOS-FORNOS**

Orientador: _____
Prof. Dr. José Leandro Félix Salles, UFES

Banca Examinadora:

Prof. Dr. Claudio Garcia, USP
Doutor pela Universidade de São Paulo – São Paulo, Brasil

Profa. Dra. Jussara Farias Fardin, UFES
Doutor pela Universidade Estadual de Campinas – Campinas, Brasil

Prof. Dr. Alexandre Loureiros Rodrigues, UFES
Doutor pela Lancaster University – Lancaster, Inglaterra

Prof. Dr. Patrick Marques Ciarelli, UFES
Doutor pela Universidade Federal do Espírito Santo – Vitória, Brasil

Coordenador do PPGEE: _____
Prof. Dr. Moisés Renato Nunes Ribeiro

VITÓRIA-ES, Julho de 2016.

Ficha catalográfica

Folha de assinaturas

Em branco

AGRADECIMENTOS

Primeiramente, agradeço à Deus por tudo que foi colocado em minha vida e pela fé na caminhada.

Dedico esta tese à minha querida esposa Isabele, por ser minha maior incentivadora e fonte infinita de paciência quando precisava me ausentar do convívio da família para trabalhar neste projeto, às minhas duas filhas, Camila e Carolina, que são o motivo da minha persistência em melhorar como ser humano, pai e profissional e aos meus pais pelos ensinamentos e orientações repassados aos longos dos anos que me colocaram no caminho que, acredito, seja o correto.

Agradeço à ArcelorMittal Tubarão, especialmente ao Engenheiro Douglas Ruy, pelo incentivo e oportunidade no desenvolvimento desse projeto. Por fim, agradeço a todos os meus mestres, principalmente meu orientador Professor Dr. José Leandro Félix Salles, pelos ensinamentos, paciência e profissionalismo.

“Se você só tem um martelo, tenderá a ver todos os problemas como pregos.” - Abraham Maslow

RESUMO

A operação de extração de material do interior do alto-forno é realizada com significativo grau de incerteza, dentre outros motivos, pois a medição do nível dos líquidos não pode ser medido diretamente. Neste trabalho é apresentado um sistema para previsão do nível dos líquidos no cadinho do alto-forno através da medição da força-eletromotriz gerada na carcaça baseado em um modelo sazonal autoregressivo integrado e de médias móveis (*Seasonal Autoregressive Integrated Moving Average - SARIMA*). Os estudos mostraram que esta força-eletromotriz é uma série temporal não-estacionária, não-linear, apresenta um forte comportamento sazonal e que é fortemente correlacionada com o nível de líquidos. Foram realizadas algumas comparações com modelos não-lineares baseados em redes neurais artificiais com atrasos de tempo (*Time Delay Neural Networking - TDNN*) e os resultados indicam que o modelo não-linear apresenta melhor performance de previsão. Esta metodologia consiste na estratégia para a análise, identificação, filtragem e previsão do nível dos líquidos através de modelo TDNN obtendo-se ao final do processo uma previsão com precisão satisfatória. A previsão do nível dos líquidos com horizonte de até 1 hora à frente pode ajudar os operadores e engenheiros durante o controle e otimização do processo de produção de altos-fornos trazendo maior segurança e ganhos financeiros.

Palavras-chave: Alto-forno, nível, cadinho, previsão, séries temporais, TDNN, não-linear, SARIMA, algoritmo genético.

ABSTRACT

The operation of material extraction from blast furnace is carried out with a significant degree of uncertainty, among other reasons, because the measuring level of liquids cannot be measured directly. This thesis presents a system for forecasting the level of liquid in the blast furnace hearth by measuring the electromotive force generated in shell based on a model seasonal autoregressive integrated moving average (SARIMA). This work has shown electromotive force is a non-stationary and nonlinear time series with a strong seasonal behavior that is strongly correlated with the level of liquids. Some comparisons were made with models based on artificial neural networks with time delay (TDNN) and the results indicated that the nonlinear model has better forecasting performance. This methodology consists of the strategy for analysis, identification, filtering and prediction of the level of liquids through TDNN models achieving at the end of the process a prediction with satisfactory accuracy. The forecast level of liquids with horizon up to 1 hour ahead can help operators and engineers during the control and process optimization of the production of blast furnaces increasing safety and financial gains.

Keywords: Blast Furnace, Hearth, Level, Forecasting, Time Series, TDNN, Nonlinear, SARIMA, Genetic Algorithm.

LISTA DE ILUSTRAÇÕES

Figura 1:	Filtro Digital de Resposta ao Impulso Finita - FIR.	21
Figura 2:	Filtro Digital de Resposta ao Impulso Infinita - IIR.	22
Figura 3:	Especificações de Filtros FIR: (a) absoluta (b) relativa. Fonte:(INGLE, 2012)	24
Figura 4:	Especificações usuais para filtros janela de Kaiser. Fonte:(MATLAB, 2015)	26
Figura 5:	Filtro linear, com entrada a_t , saída Z_t e função de transferência $\Psi_{(z-1)}$	27
Figura 6:	Exemplos simulados de séries temporais não-estacionárias.	31
Figura 7:	Seleção das bandas de frequências dominantes. (DEPIZZOL, 2011).	34
Figura 8:	Representação do cálculo das bordas dos filtros passa-banda a partir da PSD, para um sinal com múltiplas frequências de oscilação (DEPIZZOL, 2011).	35
Figura 9:	Processo de Previsão de Séries Temporais com RNAs.	36
Figura 10:	Neurônio Biológico.	37
Figura 11:	Neurônio Artificial MCP (McCulloch e Pitts).	37
Figura 12:	Redes Neurais Artificiais - Rede <i>Feedforward</i> (a) e Rede <i>Feedback</i> (b).	38
Figura 13:	Arquitetura de uma Rede Neural com atraso de tempo.	40
Figura 14:	Arquitetura de uma Rede Neural com atraso de tempo e recorrentes.	40
Figura 15:	Representação gráfica da roleta de seleção, onde o número de alternativas m é 6. O número dentro dos arcos corresponde à probabilidade de cada alternativa ser selecionada.	45
Figura 16:	Exemplo de cruzamento binário de um ponto.	46
Figura 17:	O processo do Alto-Forno.	50
Figura 18:	Visão geral de um alto-forno.	51
Figura 19:	Redimento de um Alto-forno.	52
Figura 20:	Cadinho de um alto-forno.	53
Figura 21:	Padrão típico de operação de casa de corrida de um alto-forno.	54
Figura 22:	Circuito elétrico equivalente à geração de FEM na carcaça de um alto-forno.	55
Figura 23:	Comportamento da FEM medida em relação ao esgotamento dos Líquidos no cadinho.	57
Figura 24:	Identificação de problemas de esgotamento do cadinho pela FEM.	58
Figura 25:	Disposição Física dos Sensores FEM do Alto-Forno 1 (Vista Superior e de Perfil).	59
Figura 26:	Comportamento da FEM medida na carcaça do Alto-Forno 1.	60
Figura 27:	Tela do Nível dos Líquidos FEM do Alto-forno 1.	61
Figura 28:	Análise Temporal da FEM.	62

Figura 29:	Transformada de Fourier da FEM e do Nível de Líquidos.	62
Figura 30:	Análise Temporal do Nível de Líquidos.	63
Figura 31:	Visão panorâmica do Alto-Forno #1 da ArcelorMittal Tubarão.	64
Figura 32:	Comparação entre médias e variâncias de cinco pacotes de dados do Nível.	66
Figura 33:	Sinal Filtrado de Nível dos Líquidos quanto à Linearidade.	66
Figura 34:	Autocorrelação e Autocorrelação Parcial do Nível de Líquidos.	67
Figura 35:	Autocorrelação e Autocorrelação Parcial do Nível de Líquidos Diferenciado.	68
Figura 36:	Comparação entre médias e variâncias de cinco pacotes de dados do Nível diferenciado.	69
Figura 37:	Sinal de Nível Diferenciado após Transformação de BoxCox.	69
Figura 38:	Previsão um Passo à Frente com Modelo SARIMA Adaptativo	73
Figura 39:	Auto-correlação dos Resíduos do Modelo SARIMA para Previsão de um Passo à Frente.	74
Figura 40:	Correlação Cruzada entre os Resíduos do Modelo SARIMA para Previsão de um Passo à Frente e os Valores Passados de Z_k	74
Figura 41:	Rede Neural Artificial com Atraso.	75
Figura 42:	Rede Neural Artificial Recorrente.	75
Figura 43:	Comparativo entre Modelos - MSE e MAE.	76
Figura 44:	Previsão 1 Passo à Frente - SARIMA x TDNN.	77
Figura 45:	Previsão 8 Passos à Frente - SARIMA x TDNN.	77
Figura 46:	Previsão 12 Passos à Frente - SARIMA x TDNN.	78
Figura 47:	Interface entre MELP e o Sistema de Controle do Alto-forno 1.	79
Figura 48:	Análise do tamanho de pacotes de dados versus número de iterações.	79
Figura 49:	Simulação do MELP para 1 hora à frente com 100 iterações por ciclo.	80
Figura 50:	Preditor <i>On-line</i> do Nível de Líquidos para 12 passos à frente.	80
Figura 51:	Diagrama da Filtragem do sinal FEM	82
Figura 52:	Transformada de Fourier do sinal FEM com filtro FIR.	82
Figura 53:	Distribuição normal do FEM filtrado.	83
Figura 54:	Autocorrelação e autocorrelação parcial do sinal FEM filtrado.	84
Figura 55:	Comportamento do novo sinal filtrado e diferenciado.	85
Figura 56:	Modelos ARMA com Parâmetros Estimados por algoritmos SQP e DE.	87
Figura 57:	Modelos SARIMA com Parâmetros Estimados por algoritmos SQP e DE.	88
Figura 58:	TDNN - Atrasos x Erros de Previsão.	89
Figura 59:	Comparação entre Modelos de Previsão - MSE.	90
Figura 60:	Comparação entre Modelos de Previsão - MAE.	91
Figura 61:	Comparação entre Modelos de Previsão - Análise Temporal para 12 Passos à Frente.	92
Figura 62:	Comparação entre Modelos de Previsão - Erros para 12 Passos à Frente.	92
Figura 63:	Diagrama de Previsão do Sinal EMF	93
Figura 64:	Previsão on-line do MELP para 12 passos à frente.	94
Figura 65:	Erros de previsão on-line do MELP para 12 passos à frente.	94

LISTA DE TABELAS

Tabela 1:	Teste de Estacionariedade - Sinal Filtrado versus FEM	65
Tabela 2:	Teste de Linearidade - Nível de Líquidos	65
Tabela 3:	Teste de Linearidade - Nível de Líquidos Diferenciado	68
Tabela 4:	Modelos SARIMA(p,0,q)(P,1,Q)	70
Tabela 5:	Comparativo entre Modelos - MSE	75
Tabela 6:	Comparativo entre Modelos - MAE	76
Tabela 7:	Teste de Estacionariedade - FEM x FEM Filtrado	83
Tabela 8:	Teste de Linearidade - FEM Filtrado	83
Tabela 9:	Detecção de oscilação no FEM filtrado	84
Tabela 10:	Teste de linearidade - Sinal filtrado e diferenciado	84
Tabela 11:	IMSE(12) e IMAE(12) para ARMA(p, 0, q)	86
Tabela 12:	IMSE(12) e IMAE(12) para SARIMA(p, 0, q)(0, 1, 0) ₃₂	86
Tabela 13:	IMSE(12) e IMAE(12) para SARIMA(3, 0, 0)(P, 1, Q) ₃₂	87
Tabela 14:	Modelos TDNN (neurônios,atrasos)	89

LISTA DE ABREVIATURAS

ACF	<i>Autocorrelation Function</i>
AELP	<i>Adaptive EMF Level Predictor</i>
AIC	<i>Akaike Information Criterion</i>
AF	Alto-forno
ANN	<i>Artificial Neural Networks</i>
BIC	<i>Bayesian Information Criterion</i>
BFG	<i>Blast Furnace Gas</i>
DCS	<i>Digital Control System</i>
DSP	<i>Digital Signal Processing</i>
EMF	<i>Electromotive Force</i>
FEM	Força Eletromotriz
FIR	<i>Finite Impulse Response</i>
MAE	<i>Mean Absolute Error</i>
MELP	<i>Multi-step EMF Level Predictor</i>
MSE	<i>Mean Squared Error</i>
IIR	<i>Infinite Impulse Response</i>
ODC	<i>Oscillation, Detection and Characterization</i>
OPC	<i>OLE for Process Control</i>
PACF	<i>Partial Autocorrelation Function</i>
RNN	<i>Recurrent Neural Networks</i>
SARIMA	<i>Seasonal Autoregressive Integrated Moving Average</i>
SDCD	Sistema Digital de Controle Distribuido
SQP	<i>Sequential Quadratic Programming</i>
TDNN	<i>Time Delay Neural Networks</i>
NLP	<i>Non-Linear Optimization Problems</i>
QP	<i>Quadratic Programming</i>

SUMÁRIO

1	INTRODUÇÃO	16
1.1	Objetivos do Trabalho	18
1.2	Organização da Tese	18
1.3	Publicação do Trabalho	18
2	SÉRIES TEMPORAIS: IDENTIFICAÇÃO E ESTIMAÇÃO DE PARÂMETROS	20
2.1	Pré-tratamento do Sinal	21
2.2	Séries Temporais Lineares	26
2.2.1	Modelos Autoregressivos (AR)	28
2.2.2	Modelos Médias Móveis (MA)	28
2.2.3	Modelos Integradores (I)	29
2.2.4	Modelos Sazonais (S)	29
2.2.5	Identificação de Modelos SARIMA	29
2.2.6	Testes de Linearidade	30
2.3	Séries Temporais Não-Estacionárias	31
2.3.1	Séries Trend-Estacionárias	31
2.3.2	Séries com Mudanças de Nível	32
2.3.3	Séries com Mudanças na Variância	32
2.3.4	Séries com Raízes Unitárias	32
2.3.5	Testes de Estacionariedade	32
2.3.6	Deteção e Caracterização de Oscilações	32
2.4	Séries Temporais Usando Redes Neurais Artificiais	35
2.4.1	Previsão de Séries Temporais	36
2.4.2	Uma Abordagem Biológica	36
2.4.3	Arquitetura	38
2.4.4	Treinamento	39
2.4.5	Redes Neurais com Atraso de Tempo	39
2.4.6	Redes Neurais Recorrentes	40
2.5	Estimação de Parâmetros de Séries Temporais	41
2.5.1	Otimização sem restrições	41
2.5.2	Otimização com restrições	42
2.5.3	Programação Sequencial Quadrática	43
2.5.4	Algoritmos Genéticos	44

3	NÍVEL DE LÍQUIDOS EM CADINHO DE ALTOS-FORNOS	49
3.1	Produção de Ferro-Gusa em Alto-Forno	49
3.2	O Cadinho do Alto-Forno	52
3.3	Força-Eletromotriz Gerada na Carcaça do Alto-Forno	54
3.4	A Relação entre o Processo de Esgotamento do Cadinho e a FEM	58
3.5	Instalação do Sistema de Medição do Sinal FEM	59
3.6	Filtragem do Sinal FEM	61
4	MODELAGEM E PREVISÃO DO NÍVEL DE LÍQUIDOS	64
4.1	Filtragem através do Sistema Digital de Controle	65
4.2	Modelagem e Previsão do Nível de Líquidos	69
4.2.1	Previsão com Modelos SARIMA Adaptativos	69
4.2.2	Previsão com Redes Neurais	72
4.2.3	Comparativo entre Modelos	75
4.2.4	Validação do Modelo	76
4.3	Filtragem através do MATLAB	80
4.4	Caracterização do Sinal após Nova Filtragem	82
4.5	Modelagem e Previsão após Nova Filtragem	85
4.5.1	Aproximação Linear	85
4.5.2	Estimação de Parâmetros do Modelo SARIMA com Algoritmo Genético	87
4.5.3	Modelos com Redes Neurais	89
4.5.4	Comparação entre os Modelos	90
4.6	Validação do Modelo em Tempo Real	93
5	CONCLUSÃO	95
5.1	Recomendações para Trabalhos Futuros	96
	REFERÊNCIAS	97
	APÊNDICE A REAÇÕES QUÍMICAS NO ALTO-FORNO	103
	APÊNDICE B DEFINIÇÕES GERAIS EM ESTATÍSTICA	106
B.1	Esperança de Processos Estocásticos	106
B.2	Autocovariância	107
B.3	Autocorrelação	107
B.4	Autocorrelação Parcial	107
B.5	Estacionaridade	107
B.6	Normalidade	108
B.7	Ruído Branco	108
B.8	Densidade Espectral de Potência	108

1 INTRODUÇÃO

Existem duas alternativas possíveis para a fabricação de produtos de aço: a produção de ferro-gusa no alto-forno com posterior refinamento na aciaria ou o processo de produção em aciarias de fornos elétricos tendo como base a sucata metálica. Ambas as alternativas diferem, respectivamente, quanto ao tipo de produtos obtidos bem como a matéria prima utilizada. O alto-forno é a unidade principal para produção de ferro-gusa para uma usina siderúrgica integrada. No processo de alto-forno, o minério-de-ferro e os combustíveis redutores (carvão ou coque) são transformados em ferro-gusa, ou simplesmente gusa, escória e gás de alto-forno (*blast furnace gas*- BFG). O ferro-gusa (ou gusa) se forma pela redução do minério de ferro, enquanto a escória se forma a partir da aglomeração dos minerais e das cinzas provenientes do coque ou carvão. O gusa e a escória não se misturam, ocupando áreas separadas no fundo do alto-forno, conhecido por cadinho, pois possuem densidades distintas (GEERDES et al., 2007).

As matérias-primas do gusa (sinter, coque, minério e pelota) são carregadas pela parte superior do alto-forno, enquanto que o combustível adicional é injetado pelas ventaneiras. A carga desce regularmente devido à pequena granulometria da matéria-prima que faz com que o gás flua uniformemente, aumentando o contato com os sólidos em descida. Durante a descida, parte da carga de minério-de-ferro reage com os gases redutores que sobem, reduzindo-se parcialmente os óxidos de ferro. Quando a temperatura é alta o bastante, o ferro (óxido) inicia o amolecimento e eventualmente se funde. O fundido (ferro e escória) e o coque remanescente se estabelecem no cadinho por onde são extraídos pelos furos nas paredes do cadinho chamados de furos-de-gusa. O coque remanescente forma uma camada porosa no fundo do cadinho denominada de homem-morto (*deadman*). Por perfurações nos furos-de-gusa, o cadinho é esvaziado. O metal quente líquido é então vazado em carros-torpedo e transferido para a aciaria (próxima planta no processo siderúrgico) onde se transformará em diversos tipos de aço (POST, 2003).

A condição operacional do cadinho (nível dos líquidos) é muito importante pois afeta diretamente as características de drenagem do alto-forno e, também, a qualidade do produto. O processo de esgotamento do cadinho envolve transporte de massa e energia, afetando a distribuição de fluxo gasoso dentro do alto-forno. A operação ideal da área de extração dos líquidos, conhecida por casa-de-corrída, para um alto-forno grande e de alta produtividade, é um esgotamento contínuo e com ciclos similares entre os furos-de-gusa, resultando em um fluxo de escória quase contínuo (MIELENZ; KRUNER; KOCHNER, 2008). Isto demanda um planejamento e execução precisos para as atividades de abertura e fechamento dos furos-de-gusa.

Devido à erosão, os furos-de-gusa aumentam de tamanho durante o esgotamento e conseqüentemente, a vazão de saída dos líquidos vai aumentando ao longo do esgotamento, fazendo-se necessário o fechamento do furo-de-gusa de tempos em tempos para

prevenir o contato direto entre o ar soprado pela ventaneiras e os furos-de-gusa, conhecido como sopro. O processo de abertura e fechamento dos furos-de-gusa afeta a vida-útil dos tijolos refratários desta região já que envolve impactos e vibrações significativas nas paredes do alto-forno (NIGHTINGALE; TANZIL, 2001).

O processo de esgotamento do cadinho do alto-forno é usualmente realizado com base na experiência operacional e envolve um elevado grau de incertezas devido à subjetividade envolvida, pois não há medições diretas dos níveis de ferro-gusa e escória no cadinho. Contudo, recentemente, alguns estudos foram realizados no sentido de fazer esta operação ser mais confiável (POOS, 1998; THIEMANN, 2001; FRANSSSEN, 2003; PETERS, 2002; RUTHER, 2003; TORRKULLA, 2002; TORRKULLA; SAXEN, 2000; BRANNBACKA; SAXEN, 2001).

Este dado é um importante parâmetro operacional pois, além dos ganhos diretos de estabilidade e segurança operacionais, envolve a possibilidade de otimização de processos e, conseqüentemente, a redução de custos.

Outros pesquisadores realizaram um estudo onde se desenvolveu um modelo matemático (BRANNBACKA; SAXEN, 2005) através de um modelo caixa-branca, ou seja, baseado em equações físico-químicas do processo de fluxo de líquidos dentro do cadinho, que apesar da maior precisão, demandam ajustes de constantes e um maior conhecimento sobre o processo a ser modelado. Os erros de modelagem foram corrigidos através do filtro de Kalman estendido associado com a força eletromotriz da carcaça (*electromotive force- EMF*). Este modelo foi usado para a previsão de curto prazo do momento em que o cadinho ficar vazio, com precisão de até 15 minutos. Contudo, este intervalo de tempo não é suficiente para que os operadores realizem ajustes no processo caso as condições operacionais sejam alteradas. Isto é devido à dinâmica deste processo ser muito lenta, ou seja, qualquer alteração operacional leva mais de 15 minutos para surtir efeito nas variáveis físico-químicas do alto-forno. Além disso, o ciclo de extração de líquidos do cadinho demanda um longo período (em média de 160 minutos neste estudo específico). Portanto, ao longo deste processo de extração é importante que se faça uma previsão de longo prazo para detectar os níveis crítico de operação, ou seja, quando o cadinho se encontra nos níveis máximo e mínimo.

No trabalho de (ALTER; BRUNNER; HOMES, 2013), um caso real de monitoramento do nível dos líquidos usando o FEM foi apresentado, mas nenhum modelo de previsão foi proposto. Outros pesquisadores (GOMES; SALLES, 2011) desenvolveram um modelo “caixa-preta” preliminar baseado em séries temporais não-sazonais que realiza previsão de curto prazo de até 15 minutos a frente.

Neste trabalho, alguns modelos de previsão de múltiplos passos são propostos para previsões de longo prazo do nível de líquidos no cadinho. Os primeiros testes mostraram que o sinal FEM apresenta um comportamento não-estacionário, fortes sazonalidades e não-linearidades. Após isto, dois modelos lineares foram testados, ARMA (*autoregressive moving average*) e SARIMA (*seasonal autoregressive integrated moving average*)(BOX; JENKINS; REINSEL, 1994), onde os parâmetros foram estimados usando-se algoritmos genéticos e estimadores quadráticos. Estes modelos foram comparados com modelos de redes neurais com atraso de tempo (*time-delay neural network-TDNN*) com base nos erros de previsão (GOMES; COCO; SALLES, 2016).

O modelo TDNN representou adequadamente o comportamento do nível dos líquidos no interior do cadinho mesmo quando as condições operacionais de esgotamento mudam, permitindo a previsão do nível para até uma hora à frente com boa precisão. Este horizonte de previsão é suficientemente longo para permitir que os operadores do alto-

forno tenham tempo suficiente para atuar em caso de instabilidades ou ajustar o processo de esgotamento antecipadamente, se necessário. Um estudo de caso foi desenvolvido e apresentado aplicando-se esta metodologia para um alto-forno ativo localizado no Brasil. Este trabalho resultou em uma solicitação de patente junto ao Instituto Nacional de Propriedade Intelectual (INPI) que está em processo de análise técnica.

1.1 Objetivos do Trabalho

O objetivo geral deste trabalho é a implantação de um sistema de medição e previsão do nível dos líquidos no interior do alto-forno com base em modelos caixa-preta (modelagem que não utiliza equações físico-químicas do processo) pois é premissa a busca por uma metodologia que possa ser utilizada em altos-fornos sistemicamente e sem a necessidade de profundos conhecimentos a respeito do processo e suas nuances, como por exemplo, as complexas reações que ocorrem no interior da planta.

Os objetivos específicos desta pesquisa são:

- Implementar sistema de comunicação entre o sistema de previsão de nível de líquidos do alto-forno usando o software matemático Matlab para aquisição de dados em tempo real;
- Aperfeiçoar o modelo autoregressivo gerado em (GOMES, 2010) visando menor erro de previsão utilizando-se ferramentas como séries temporais sazonais (modelos SARIMA), algoritmos genéticos e redes neurais;
- Gerar previsões de até uma hora à frente para o nível dos líquidos no cadinho do alto-forno com performance satisfatória, mesmo durante instabilidades operacionais.

Além da relevância do tema, trata-se de um sistema que, quando operacional, será inédito em altos-fornos industriais.

1.2 Organização da Tese

O texto está organizado da seguinte forma: no Capítulo 2, a análise de séries temporais é apresentada como ferramenta para a modelagem e previsão do nível dos líquidos no cadinho do alto-forno. No Capítulo 3 é realizada uma descrição sucinta do processo de produção do ferro-gusa e o fenômeno gerador da força-eletromotriz na carcaça e sua relação com o processo. No Capítulo 4 é apresentada a descrição detalhada de todo o desenvolvimento do sistema de medição de nível no processo produtivo, assim como os resultados obtidos em cada uma das etapas do trabalho. Finalmente, no Capítulo 5 são apresentadas as conclusões do trabalho e também alguns possíveis caminhos para continuação das análises e desenvolvimentos para trabalhos futuros.

1.3 Publicação do Trabalho

Em conformidade com a Resolução PPGEE da UFES de número 02/2010 Artigo 3, ratificamos que este trabalho foi publicado em periódico científico intitulado *IEEE - Transactions on Automation Science and Engineering*. O trabalho foi publicado em 28 de Março de 2016 com o título *Multistep Forecasting Models of the Liquid Level*

in a Blast Furnace Hearth (Volume:PP, Issue:99),ISSN:1545 – 5955, DOI:10.1109 / TASE.2016.2538560 (GOMES; COCO; SALLES, 2016) habilitando o candidato à sua defesa pública.

2 SÉRIES TEMPORAIS: IDENTIFICAÇÃO E ESTIMAÇÃO DE PARÂMETROS

As séries temporais sempre tiveram um importante papel nas ciências naturais. A astronomia babilônica utilizava séries temporais nas posições relativas das estrelas e planetas para prever eventos astronômicos. Observações dos movimentos dos planetas formaram a base das leis que Johannes Kepler formulou (G.KIRCHGASSNER; WOLTERS, 2007). Séries temporais podem ser definidas como pacotes de observações quantitativas arranjadas em ordem cronológica em que o tempo é considerado uma variável discreta.

A análise de séries temporais permite a detecção de regularidades nas observações das variáveis e encontrar a lei que as define. Através desta análise também é possível prever observações futuras com precisão. A idéia básica por trás dos procedimentos, que vem desde os babilônicos, é que é possível decompor as séries temporais em números finitos de componentes independentes que juntas apresentam regularidade e podem ser calculadas antecipadamente. Neste procedimento é necessário que haja diferentes fatores independentes impactando na variável.

A decomposição das séries temporais em componentes não observáveis dependentes de diferentes fatores causais foi utilizada sistematicamente na análise de séries temporais desenvolvida por Warren M. Persons em 1919 (G.KIRCHGASSNER; WOLTERS, 2007). Ele identificou e separou as séries temporais em quatro componentes: componente de longo prazo, ou *trend*; componente cíclica com períodos de mais de um ano, ou ciclo de negócios; componente que contém subidas e descidas no ano, ou ciclo sazonal e componente que contém comportamentos que não se enquadram nas demais, ou resíduo.

Na década de 1920, uma visão totalmente diferente foi amplamente difundida utilizando-se de análises estatísticas das séries temporais. Os procedimentos descritivos da análise clássica de séries temporais foi abandonado e, em seu lugar, surgiram resultados provenientes de métodos de teorias probabilísticas e estatísticas. Destes, surgiram as regras de processos estocásticos que seriam utilizadas amplamente na análise de séries temporais. Esta nova metodologia de análise de séries temporais as considerava como uma realização de um processo estocástico, focando-se nos termos não-determinísticos como estruturas complexas e dependentes. Os primeiros passos foram dados por Evgenij Evgenievich Slutsky e George Udny Yule no começo da década de 1920, que começaram a representar as séries como somas de processos estocásticos através de médias móveis e processos auto-regressivos. Esta prática foi amplamente divulgada por George E.P. Box e Gwilym M. Jenkins em 1970 que desenvolveram métodos empíricos de implementação destas regras.

Neste capítulo serão apresentados alguns aspectos importantes a respeito de séries temporais. O primeiro deles é a importância do pré-tratamento e filtragem dos sinais

analisados para que os modelos consigam sucesso na captura dos principais componentes contidos nas séries. Na sequência serão apresentadas as séries temporais lineares, não estacionárias e não lineares. Na Seção 2.4 é apresentado o uso de redes neurais como ferramentas de modelagem de séries temporais. Por fim, algumas técnicas de estimação de parâmetros de séries temporais são apresentadas na Seção 2.5. No apêndice B estão descritos diversas ferramentas e conceitos de processos estocásticos utilizados ao longo deste capítulo.

2.1 Pré-tratamento do Sinal

Filtros digitais são componentes importantes da área de processamento digital de sinais (*Digital Signal Processing - DSP*) e são utilizados, geralmente, com foco nos seguintes objetivos principais: (1) separação de sinais que estão combinados e, (2) recuperação de sinais que foram distorcidos de alguma forma. Esta última é necessária quando um sinal está contaminado com interferências, ruídos, ou outros sinais indesejados.

É comum em DSP se referir aos sinais de entrada e saída do filtro como estando no domínio do tempo. Isto acontece pois os sinais encontram-se, usualmente, amostrados em intervalos regulares no tempo. A forma mais direta de se implementar um filtro digital é pela convolução do sinal de entrada com a resposta ao impulso do filtro digital. Todos os filtros digitais lineares podem ser obtidos desta forma. Existe uma outra forma de se fazer a filtragem digital chamada recursão. Quando um filtro é implementado através de convolução, cada amostra na saída é calculada pela soma ponderada das amostras na entrada somadas. Filtros recursivos são uma extensão disto, usando valores calculados previamente da saída, além dos pontos da entrada.

Para se encontrar a resposta ao impulso de um filtro recursivo, simplesmente injeta-se um impulso, e avalia-se o que sai do filtro. As respostas ao impulso de filtros recursivos são compostos por senóides que diminuem exponencialmente em amplitude. No começo, a resposta é composta por sinais muito longos caindo abaixo dos níveis detectáveis de ruído após algum tempo. Por causa desta característica, filtros recursivos são chamados de filtros de resposta ao impulso infinita (*Infinite Impulse Response - IIR*). Em comparação, os filtros que utilizam convolução são chamados de filtros de resposta ao impulso finita (*Finite Impulse Response - FIR*) (SMITH, 1999).

Nas Figuras 1 e 2 podem-se visualizar as representações gráficas de um filtro de resposta ao impulso finita e filtro de resposta ao impulso infinita, respectivamente. Nas Equações (2.1) e (2.2) estão apresentadas suas representações matemáticas.

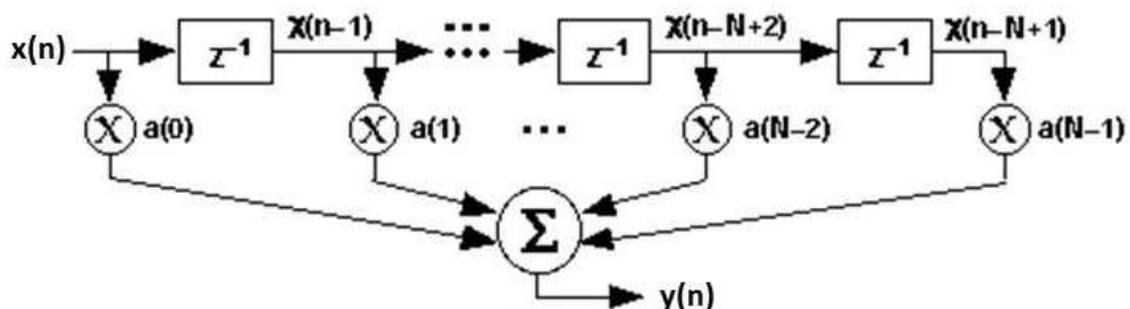


Figura 1: Filtro Digital de Resposta ao Impulso Finita - FIR.

$$\begin{aligned}
y(n)_{FIR} &= \sum_{k=0}^{N-1} a(k)x(n-k) \\
&= a_0x(n) + a_1x(n-1) + \dots + a_{N-1}x(n-N+1)
\end{aligned}
\tag{2.1}$$

onde y_n é a saída do filtro digital, x_n é o sinal a ser filtrado, n é o instante de tempo discreto, z^{-1} é o atraso discreto de primeira ordem, $a(k)$ são os coeficientes do filtro, $N-1$ é a ordem do filtro e N é o tamanho da amostra.

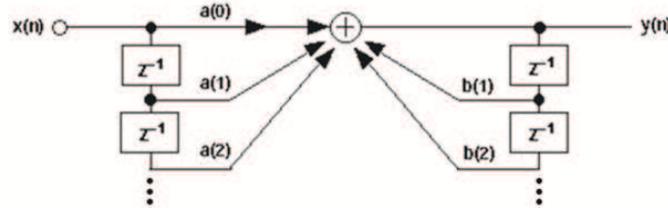


Figura 2: Filtro Digital de Resposta ao Impulso Infinita - IIR.

$$\begin{aligned}
y(n)_{IIR} &= \sum_{k=0}^{N-1} a(k)x(n-k) + \sum_{k=1}^M b(k)y(n-k) \\
&= a_0x(n) + a_1x(n-1) + \dots + a_{N-1}x(n-N+1) \\
&\quad b_1y(n-1) + b_2y(n-2) + \dots + b_My(n-M),
\end{aligned}
\tag{2.2}$$

onde y_n é a saída do filtro digital, x_n é o sinal a ser filtrado, n é o instante de tempo discreto, z^{-1} é o atraso discreto de primeira ordem, $a(k)$ são os coeficientes do filtro, $b(k)$ são os coeficientes recursivos do filtro, M é a ordem do filtro e N é o tamanho da amostra.

Usando a transformada de Fourier discreta nas Equações (2.1) e (2.2) teremos:

$$Y(z)_{FIR} = a_0X(z) + a_1z^{-1}X(z) + \dots a_kz^{-(N-1)}X(z) \tag{2.3}$$

$$\begin{aligned}
Y(z)_{IIR} &= a_0X(z) + a_1z^{-1}X(z) + \dots a_kz^{-(N-1)}X(z) + \\
&\quad + b_1z^{-1}Y(z) + b_2z^{-2}Y(z) + \dots + b_Mz^{-M}Y(z)
\end{aligned}
\tag{2.4}$$

Se dividirmos as equações (2.3) e (2.4) por $X(z)$, teremos as funções de transferência dos filtros FIR e IIR, respectivamente:

$$H(z)_{FIR} = \frac{Y(z)_{FIR}}{X(z)} = a_0 + a_1z^{-1} + \dots a_kz^{-(N-1)} \tag{2.5}$$

$$H(z)_{IIR} = \frac{Y(z)_{IIR}}{X(z)} = \frac{a_0 + a_1z^{-1} + \dots a_kz^{-(N-1)}}{1 - b_1z^{-1} - b_2z^{-2} - \dots - b_Mz^{-M}} \tag{2.6}$$

Os filtros FIR possuem algumas vantagens em relação aos filtros IIR, são elas (SHE-NOI, 2006):

- Podemos facilmente projetar um filtro FIR que atenda aos requisitos de resposta de amplitude com atraso de grupo constante. Atraso de grupo é definido como $\tau = -(d\theta/d\omega)$, onde θ é a resposta de fase do filtro. A resposta de fase do filtro com um atraso de grupo constante é uma função linear da frequência, ou seja,

transmite todas as frequências com o mesmo atraso. Em outras palavras, não haverá distorção de fase e o sinal de entrada será atrasado por uma constante de tempo quando transmitido para a saída. Um filtro com atraso de grupo constante é fortemente desejado em transmissões de sinais digitais;

- As amostras de sua resposta ao impulso unitário são os mesmos que os coeficientes da função de transferência, não havendo necessidade de calcular $h(n)$ a partir de $H(z^{-1})$ durante todas as fases do procedimento de otimização iterativa ou para desenhar as estruturas a partir de $H(z^{-1})$;
- Os filtros FIR são sempre estáveis independentes da resolução da entrada;
- Os efeitos da resolução finita na resposta específica ou a resposta no domínio do tempo ou ruído de saída é menor que os apresentados pelos filtros IIR;
- Mesmo a resposta ao impulso unitário $h(n)$ de um filtro IIR sendo uma seqüência infinitamente longa, é razoável supor na maioria dos casos práticos, que o tamanho das amostras do sinal de entrada torna-se quase desprezível após certa quantidade; assim, a escolha de uma seqüência de comprimento finito para o sinal de tempo discreto, permite a utilização de métodos numéricos poderosos para o processamento de sinais de comprimento finito.

O projeto de um filtro digital é executado em três passos:

Especificação Antes de projetar um filtro, é preciso ter em mãos algumas especificações e estas são determinadas pela aplicação, tais como, frequências de corte, *ripple*, ganho, defasagem, etc;

Aproximação Após definição das especificações, utiliza-se vários conceitos e definições para escolher um filtro que possua a descrição que se aproxime dos parâmetros especificados. Este passo é um tópico de estudo em projetos de filtros, onde o conhecimento de vários filtros facilita a definição da melhor opção para resolver cada tipo de problema;

Implementação O produto do passo acima é a descrição do filtro na forma de uma equação diferencial, uma função sistêmica $H(z)$ ou resposta ao impulso $h(n)$. A partir desta descrição, o filtro é implementado em hardware ou através de software.

Em muitas aplicações, como processamento de sinais de voz ou áudio, os filtros digitais são usados para implementar operações de seletividade de frequências. Desta forma, especificações são requeridas no domínio da frequência em termos de amplitude e resposta de fase desejadas para o filtro. Em geral, uma resposta de fase linear na banda passante do filtro é desejada. No caso de filtros FIR, esta linearidade é plenamente alcançada, mas nos filtros IIR isto não é possível. Consequentemente, após seleção do tipo de filtro (FIR ou IIR), apenas especificações de amplitude são consideradas. Estas especificações são analisadas de duas formas. A primeira abordagem é chamada especificação absoluta, e define uma série de requisitos para a função amplitude de resposta $|H(e^{j\omega})|$. Estas especificações são geralmente usadas para filtros FIR. Filtros IIR são especificados de forma um pouco diferente. A segunda abordagem é chamada de especificação relativa, definindo os requisitos em decibéis (dB), dados por:

$$dB \text{ scale} = -20 \log_{10} \frac{|H(e^{j\omega})|}{|H(e^{j\omega})|_{max}}$$

Esta especificação é mais popular na prática e é utilizada em filtros FIR e IIR (INGLE, 2012). Uma típica especificação absoluta de um filtro digital passa-baixas é mostrada na Figura 3(a), onde:

- Banda $[0, \omega_p]$ é chamada de banda passante e δ_1 é a tolerância, ou *ripple*, que aceitamos como a resposta ideal da banda passante;
- Banda $[\omega_s, \pi]$ é chamada de corte, e δ_2 é a tolerância de oscilação correspondente, ou *ripple*;
- Banda $[\omega_p, \omega_s]$ é chamada de banda de transição, não havendo restrições na resposta em amplitude desta banda.

Na Figura 3(b) pode-se visualizar especificações relativas para um filtro passa-baixa:

- R_p é a oscilação da banda passante em decibéis (*dB*);
- A_s é a atenuação da banda de corte, em decibéis (*dB*).

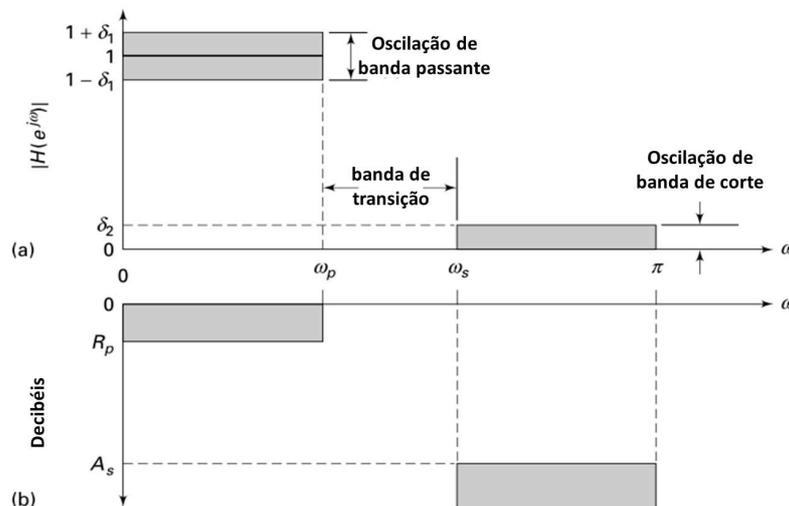


Figura 3: Especificações de Filtros FIR: (a) absoluta (b) relativa. Fonte:(INGLE, 2012)

Os parâmetros dados nestas duas classes de especificações são obviamente relacionadas. No caso de $|H(e^{j\omega})|_{max}$ das especificações absolutas igual a $(1 + \delta_1)$, teremos:

$$dB \text{ scale} = -20 \log_{10} \frac{(1 - \delta_1)}{(1 + \delta_1)} > 0 (\approx 0)$$

e

$$dB \text{ scale} = -20 \log_{10} \frac{\delta_2}{(1 + \delta_1)} > 0 (\gg 1)$$

Estas especificações foram dadas para filtros passa-baixa. Especificações similares podem ser usadas para outros tipos de filtros de seleção de frequência como passa-altas ou passa-faixa.

O método da janela (projeto de transformada de Fourier com funções de janela) é utilizado para mitigar oscilações indesejadas nas bandas passantes e de corte em filtros FIR originadas do corte abrupto na sequência de coeficientes de comprimento infinito (conhecidas como oscilações de Gibbs). É natural a busca por uma função janela que seja simétrica e que possa gradualmente ponderar os coeficientes até zero nas bordas do filtro para o comprimento de $-M \leq n \leq M$. Aplicando a sequência de janela para os coeficientes dos filtros, temos:

$$h_w(n) = h(n)w(n)$$

onde $w(n)$ é a função janela. Algumas funções janela comumente usadas em projetos de filtros FIR são (TAN; JIANG, 2013):

1. Janela retangular

$$w_{ret}(n) = 1, -M \leq n \leq M$$

2. Janela triangular (Barlett)

$$w_{tri}(n) = 1 - \frac{|n|}{M}, -M \leq n \leq M$$

3. Janela de Hanning

$$w_{han}(n) = 0.5 + 0.5 \cos\left(\frac{n\pi}{M}\right), -M \leq n \leq M$$

4. Janela de Hamming

$$w_{ham}(n) = 0.54 + 0.46 \cos\left(\frac{n\pi}{M}\right), -M \leq n \leq M$$

5. Janela de Blackman

$$w_{black}(n) = 0.42 + 0.5 \cos\left(\frac{n\pi}{M}\right) + 0.08 \cos\left(\frac{2n\pi}{M}\right), -M \leq n \leq M$$

Além destes, existe uma outra função janela bastante popular chamada janela de Kaiser. Esta função é uma janela ajustável criada por J.F. Kaiser (KAISER; SCHAFER, 1980) e definida como:

$$w_{kaiser}(n) = \frac{I_0\beta \left[\sqrt{1 - \left(1 - \frac{2n}{M-1}\right)^2} \right]}{I_0[\beta]}, \quad 0 \leq n \leq M-1 \quad (2.7)$$

onde $I_0(\cdot)$ é a função Bessel de ordem zero modificada e é descrita como sendo:

$$I_0(x) = 1 + \sum_{k=0}^{\infty} \left[\frac{(x/2)^k}{k!} \right]^2$$

que é positivo para todos os valores reais de x . O parâmetro β controla a atenuação mínima de banda de corte (A_{stop}). Na Figura 4 estão mostrados os parâmetros usuais de especificação para um filtro passa-faixa baseado em filtro de Kaiser, onde:

- A_{stop1} é a atenuação mínima da banda de corte para as baixas frequências (dB);

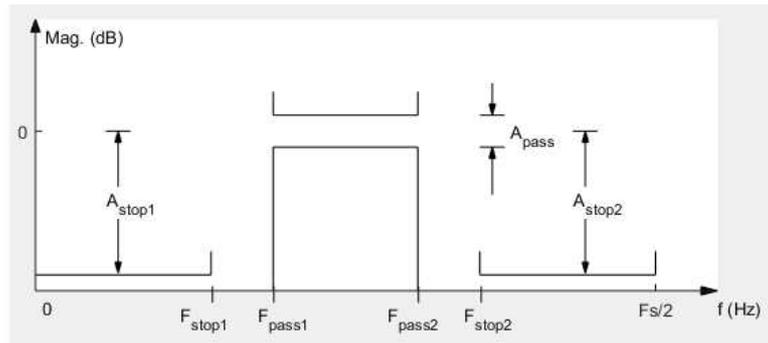


Figura 4: Especificações usuais para filtros janela de Kaiser. Fonte:(MATLAB, 2015)

- A_{stop2} é a atenuação mínima da banda de corte para as altas frequências (dB);
- F_{stop1} é a frequência inicial de corte das baixas frequências (Hz);
- F_{stop2} é a frequência inicial de corte das altas frequências (Hz);
- F_{pass1} é a mínima frequência da banda passante (Hz);
- F_{pass2} é a máxima frequência da banda passante (Hz);
- A_{pass} é a amplitude da banda passante (dB);
- F_s é a frequência de amostragem do sinal que será filtrado (Hz).

Com base nestes dados, softwares matemáticos são capazes de calcular os parâmetros contidos na Equação (2.7).

2.2 Séries Temporais Lineares

A criação de modelos para séries temporais depende de diversos fatores, tais como o comportamento do fenômeno ou o conhecimento a priori que se tem de sua natureza e do objetivo da análise. Na prática, depende também da existência de métodos apropriados de estimação e da disponibilização de softwares adequados (MORETTIN; TOLOI, 2006).

Pode-se classificar os modelos para séries temporais em duas classes: modelos paramétricos (número finito de parâmetros); modelos não-paramétricos (número infinito de parâmetros).

Na classe de modelos paramétricos, a análise é feita no domínio do tempo. Dentre estes, os mais usados são os modelos de erro (ou de regressão), os modelos auto-regressivos e de médias móveis (ARMA), os modelos auto-regressivos integrados e de médias móveis (ARIMA), modelos sazonais, modelos de memória longa, modelos estruturais e modelos não-lineares.

Os modelos não-paramétricos mais utilizados são a função de autocovariância (ou autocorrelação) e sua transformada de Fourier, ou espectro de frequência.

É importante ressaltar que, mesmo trabalhando-se com modelos paramétricos, a análise do espectro de frequência desempenha importante papel nos estudos de resposta em frequências e na área de otimização de desempenho de processos industriais, especialmente em modelos SARIMA. Estes modelos descrevem, de maneira adequada, as seguintes classes de processos (MORETTIN; TOLOI, 2006):

1. Processos lineares estacionários;

$$Z_t - \mu = \sum_{k=0}^{\infty} \varphi_k a_{t-k}, \quad \varphi_0 = 1 \quad (2.8)$$

onde a_t é ruído branco, $\mu = E(Z_t)$ é a esperança de Z_t e $\varphi_1, \varphi_2, \dots$ é uma sequência de parâmetros tal que:

$$\sum_{k=0}^{\infty} \varphi_k^2 < \infty. \quad (2.9)$$

2. Processos lineares não-estacionários homogêneos. São processos que possuem erros correlacionados e que as séries apresentam não-estacionaridade de nível ou inclinação (comportamento de um ação na bolsa de valores, por exemplo) e podem ser transformados no caso 1 por meio de diferenciação.

Os modelos analisados neste trabalho baseiam-se no estudo de modelos de filtro linear. Este modelo supõe que a série temporal é gerada por um sistema linear tendo um ruído branco como entrada como pode ser visualizado na Figura 5:

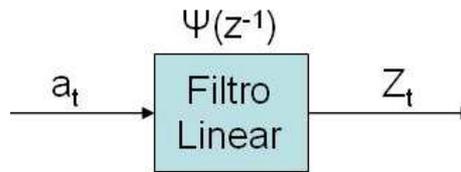


Figura 5: Filtro linear, com entrada a_t , saída Z_t e função de transferência $\Psi(z^{-1})$.

onde a_t é o ruído branco, Z_t é a série temporal e $\Psi(z^{-1})$ é denominada função de transferência do processo. Na modelagem proposta em (BOX; JENKINS; REINSEL, 1994), o filtro linear $\Psi(z^{-1})$ é definido por:

$$\Psi(z^{-1}) = \frac{\theta_q(z^{-1})}{\phi_q(z^{-1})} \quad (2.10)$$

Desta forma, os primeiros modelos Box&Jenkins são dados por:

$$\phi_q(z^{-1})Z_t = \theta_q(z^{-1})a_t \quad (2.11)$$

onde:

- $\phi(B)$ e $\theta(B)$ são polinômios do tipo $P(z^{-1}) = 1 - C_1z^{-1} - C_2z^{-2} - \dots - C_kz^{-k}$ de graus p e q , respectivamente.
- z^{-1} ou B é o operador atraso e seu efeito sobre a variável discreta é $B^k Z_t = Z_{t-k}$.

Derivado deste conceito, Box&Jenkins (BOX; JENKINS; REINSEL, 1994) desenvolveram uma classe de modelos para identificação de séries temporais lineares com sazonalidade conhecida como modelos $SARIMA(p, d, q)$ (P, D, Q)_S (*seasonal autoregressive integrated moving average*). Estes modelos SARIMA descrevem adequadamente processos lineares estacionários como mostrado em (WEI, 2007). A nomenclatura matemática geral utilizada para este tipo de classe de modelo é dada na Equação (2.12):

$$\Phi(B^S)\phi(B)\Delta_S^D\Delta^d Z_t = \Theta(B^S)\theta(B)a_t, \quad (2.12)$$

onde:

- B é o operador atraso;
- S é o período da sazonalidade;
- h é o horizonte de previsão;
- a_t é o ruído branco no tempo t ;
- Z_t é o valor do sinal no tempo t ;
- $\Delta_S^D = (1 - B^S - B^{2S}, \dots, -B^{DS})$ são os operadores das diferenças sazonais, (ordem D);
- $\Delta^d = (1 - B - B^2, \dots, -B^d)$ são os operadores das diferenças, (ordem d);
- $\Phi(B^S) = (1 - \Phi_1 B^S - \Phi_2 B^{2S}, \dots, -\Phi_P B^{PS})$ são os parâmetros sazonais autoregressivos (ordem P);
- $\phi(B) = (1 - \phi_1 B - \phi_2 B^2, \dots, -\phi_p B^p)$ são os parâmetros autoregressivos (ordem p);
- $\Theta(B^S) = (1 - \Theta_1 B^S - \Theta_2 B^{2S}, \dots, -\Theta_Q B^{QS})$ são os parâmetros sazonais médias móveis (ordem Q);
- $\theta(B) = (1 - \theta_1 B - \theta_2 B^2, \dots, -\theta_q B^q)$ são os parâmetros médias móveis (ordem q).

Deste modelo podemos extrair alguns casos particulares: modelos autoregressivos, modelos médias móveis, modelos integradores e modelos sazonais.

2.2.1 Modelos Autoregressivos (AR)

Os modelos cujos polinômios $\phi(B)$ não são iguais a 1 e $\theta(B) = 1$ são chamados modelos autoregressivos. Sendo assim, temos a seguinte equação para estes modelos:

$$\phi(B)Z_t = a_t, \quad (2.13)$$

Estes modelos são chamados autoregressivos devido ao fato de que Z_t no instante t é função apenas dos Z_s nos instantes anteriores à t .

2.2.2 Modelos Médias Móveis (MA)

Os modelos nos quais apenas os polinômios $\theta(B)$ não são iguais a 1 e $\phi(B) = 1$ são chamados modelos médias móveis. Sendo assim, temos a seguinte equação para estes modelos:

$$Z_t = \theta(B)a_t, \quad (2.14)$$

Estes modelos são chamados médias móveis devido ao fato de que Z_t é uma função soma algébrica ponderada dos a_{ts} que se movem no tempo de forma similar a uma média móvel. Contudo, isto não é verdade pois nem sempre a soma dos termos totaliza uma unidade.

2.2.3 Modelos Integradores (I)

Algumas séries não-estacionárias homogêneas é possível de haver a representação por modelos *ARMA* desde que seja realizada a diferenciação dos valores de Z_t previamente. Para estes casos, sendo Z_t não-estacionária, tomamos a diferenciação:

$$W_t = Z_t - Z_{t-1} = (1 - B)^d Z_t = \Delta^d Z_t \quad (2.15)$$

Com isto, esta série diferenciada apresenta comportamento estacionário. A quantidade de vezes que é necessário se diferenciar uma série para torná-la estacionária é chamada de ordem do modelo integrado (d) ou (D) no caso de diferenças sazonais definida abaixo:

$$W_{tD} = Z_t - Z_{t-S} = (1 - B^S)Z_t = \Delta_S Z_t \quad (2.16)$$

2.2.4 Modelos Sazonais (S)

Em alguns casos, onde Z_t exibe um comportamento sazonal determinístico com período S , pode-se utilizar o modelo abaixo:

$$Z_t = \mu_t + N_t \quad (2.17)$$

onde μ_t é uma função determinística periódica, satisfazendo $\mu_t - \mu_{t-S} = 0$, ou

$$(1 - B^S)\mu_t = 0 \quad (2.18)$$

e N_t é um processo estacionário que pode ser modelado por um *ARMA*(p, q).

De acordo com a ordem do sistema modelado, podem existir modelos mistos como, por exemplo, autoregressivos sazonais (SAR) ou não-sazonais (AR), diferenças sazonais (SI) ou não-sazonais (I) e médias móveis sazonais (SMA) ou não sazonais (MA).

2.2.5 Identificação de Modelos SARIMA

Uma metodologia bastante utilizada na análise de modelos paramétricos é conhecida como abordagem de Box&Jenkins, que foi criada em 1970 e aperfeiçoada em 1994 (BOX; JENKINS; REINSEL, 1994). Esta consiste no ajuste de modelos auto-regressivos integrados e de médias móveis, *SARIMA*(p, d, q), a um conjunto de dados (SOUZA; CA-MARCO, 2004).

A estratégia de construção do modelo é baseada no ciclo iterativo abaixo:

- (a) Uma classe de modelos é especificada para a análise;
- (b) Há uma identificação do modelo com base na análise de autocorrelações, autocorrelações parciais e outros critérios;
- (c) Os parâmetros dos modelos são estimados, usualmente por meio de mínimos quadrados;
- (d) Há a verificação ou validação do modelo encontrado, através da análise dos resíduos, para se saber se é adequado ao objetivo desejado (previsão, por exemplo).

Em caso de problemas em quaisquer uma das fases acima, todo o ciclo é repetido. A etapa crítica do processo é a identificação. Esta se divide em três partes:

- Checar a necessidade de uma transformação na série original, com o objetivo de se estabilizar sua variância (séries estacionárias). Isto pode ser feito através de análise gráfica;
- Diferenciar a série quantas vezes forem necessárias até que a série se torne estacionária de modo que o processo $\Delta^d Z_t$ seja reduzido a um $ARMA(p,q)$;
- Identificar a existência de sazonalidades no comportamento da série temporal com base na análise das frequências dominantes contidas na transformada de Fourier;
- Identificar o processo $ARMA(p,q)$ resultante, através da análise das autocorrelações e autocorrelações parciais do processo em análise comparando-as com os respectivos comportamentos dos modelos AR, MA e ARMA teóricos (MORETTIN; TOLOI, 2006).

2.2.6 Testes de Linearidade

Os modelos da classe *SARIMA* podem falhar na captura de todo o comportamento dinâmico de séries temporais reais caso estas apresentem componentes não-lineares significativos pois estes modelos são, por natureza, modelos lineares. Portanto, é importante que as análises preliminares incluam verificações de linearidade do processo a ser modelado (CORDUAS, 1994).

O teste de Lilliefors é um teste de linearidade, pela análise da distribuição normal, baseado no teste de Kolmogorov-Smirnov (KOLMOGOROV, 1933) e (SMIRNOV, 1948). Ele é usado para se testar a hipótese nula que os dados analisados são provenientes de uma população com distribuição normal, onde esta hipótese nula não especifica qual é esta distribuição normal, ou seja, não se conhece nem o valor esperado nem a variância desta distribuição (LILLIEFORS, 1967).

O teste de Kolmogorov-Smirnov é executado da seguinte forma:

- Estima-se a média e a variância da população com base no pacote de dados testado;
- Calcula-se a discrepância máxima entre a função de distribuição empírica e a função de distribuição cumulativa (*cumulative distribution function* - CDF) da distribuição normal com a média e variância estimadas;
- Avalia-se se a máxima discrepância é suficiente para ser estatisticamente significativa, impactando na rejeição da hipótese nula.

O teste de Kruskal-Wallis é um método não-paramétrico de uma via de análise de variância usado para testar se dois ou mais conjuntos de amostras provêm da mesma distribuição pela análise das variâncias. A hipótese nula testada é que todas as populações possuem funções de distribuição com médias iguais contra a hipótese alternativa de que ao menos duas das populações possuem funções de distribuição diferentes (KRUSKAL; WALLIS, 1952).

Em caso de avaliação negativa quanto à linearidade, há necessidade de utilização de ferramentas de modelagens adaptativas quanto aos parâmetros ou não-lineares como redes neurais.

2.3 Séries Temporais Não-Estacionárias

Antes de iniciar o processo de modelagem, é necessário conhecer se o sinal pode ser representado por um modelo estacionário ou se existe alguma forma de não-estacionariedade. Em caso positivo, pode ser necessário encontrar uma forma de removê-la ou, pelo menos mitigá-la para reduzir os esforços computacionais durante o processo de modelagem e então, em muitos casos, alcançar melhor precisão de previsão (HAMILTON, 1994).

As variáveis econômicas e de processos industriais possuem, em geral, algum tipo de tendência, por exemplo, elas tendem, sistematicamente, a crescer ou diminuir com o passar do tempo. O comportamento de tendência significa que o valor esperado muda com o passar do tempo, isto não estando em acordo com o significado de estacionariedade assumido anteriormente.

2.3.1 Séries Trend-Estacionárias

Em alguns casos, a tendência é tão sistemática que os desvios em torno da tendência são considerados uma variável estacionária. Neste caso, nós podemos analisar o desvio da tendência ao invés da variável original, e como esta nova variável é estacionária, o processo usual de análise de séries estacionárias é usado. Uma série temporal que flutua ao redor de uma tendência determinística linear, de uma maneira estacionária, é chamado de processo *trend*-estacionário. Como exemplo, podemos analisar um processo autoregressivo estacionário de primeira ordem AR(1) com média zero,

$$\tilde{x}_t = \theta \tilde{x}_{t-1} + \epsilon_t, \quad |\theta| < 1, \quad (2.19)$$

e um novo processo, x_t , definido como um processo estacionário somado a um termo de tendência e uma constante,

$$x_t = \theta \tilde{x}_t + \mu_0 + \mu_1 t. \quad (2.20)$$

Como \tilde{x}_t é um processo estacionário, x_t é estacionário ao redor da média da tendência, $E[x_t] = \mu_0 + \mu_1 t$, ou seja, é *trend*-estacionário. Uma realização com 200 observações do processo \tilde{x}_t e x_t (com $\theta = 0.5$) é apresentado na Figura 6 (A).

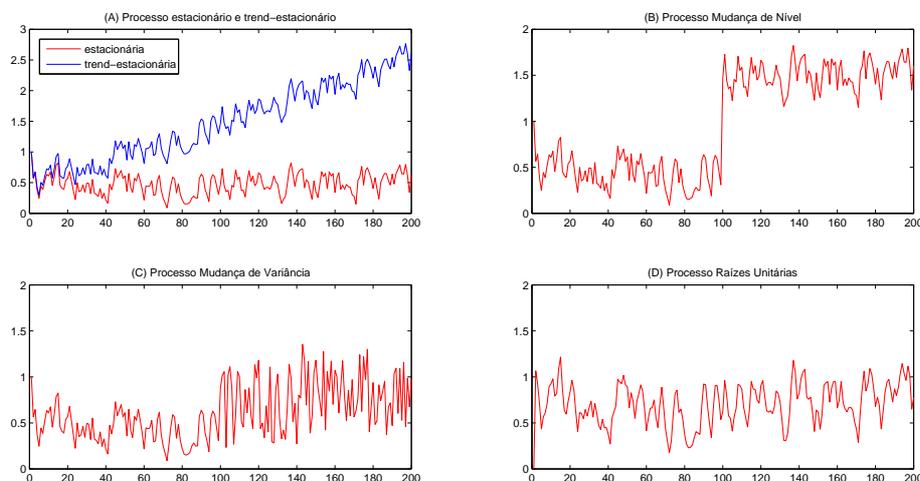


Figura 6: Exemplos simulados de séries temporais não-estacionárias.

2.3.2 Séries com Mudanças de Nível

Outro tipo de não-estacionariedade em série temporais é notada quando existem variações na média em algum ponto no tempo. Como exemplo, a média da série temporal pode ser μ_1 para a primeira metade da amostra e μ_2 para a segunda metade. Este tipo de série pode ser visualizada na Figura 6 (B). A mudança de nível pode estar associada a mudanças de estrutura ou transições entre pontos de operação. Do ponto de vista de modelagem, é usual considerar mudanças na média como determinísticas e incluir uma variável ‘burra’ no modelo regressivo. A variável ‘burra’ é definida como ($D_t = 0$, se $t \geq T_0$), sendo zero antes da observação T_0 e um após este instante, resultando na equação abaixo:

$$y_t = \beta_0 + \beta_1 x_t + \beta_2 D_t + \epsilon. \quad (2.21)$$

2.3.3 Séries com Mudanças na Variância

Um terceiro tipo de não-estacionariedade é a relacionada às mudanças na variância. A Figura 6 (C) ilustra este caso.

$$y_t = 0,5.y_{t-1} + \epsilon. \quad (2.22)$$

onde $\epsilon \sim N(0, 1)$ para $t = 1, 2, \dots, 100$, e $\epsilon \sim N(0, 5)$ para $t = 101, 102, \dots, 200$. Novamente, a interpretação nos leva a crer que a série cobre regimes de operação diferentes, onde um regime aparentemente é mais volátil que outro. No caso de pedaços de amostras muito grandes, uma solução natural é modelar os regimes separadamente. Uma solução alternativa é tentar modelar as mudanças na variância através de modelos chamados ARCH (*autoregressive conditional heteroskedasticity*).

2.3.4 Séries com Raízes Unitárias

O último tipo de não-estacionariedade mais comum é chamada de série com raízes unitárias em modelos autoregressivos. A Figura 6 (D) apresenta o caso também chamado de passeio aleatório (*random walk*),

$$y_t = y_{t-1} + \epsilon_t, \quad \epsilon_t \sim N(0, 1), \quad (2.23)$$

que possui uma raiz unitária no polinômio característico. É possível notar que o passeio aleatório não possui amarrações e varia livremente para cima e para baixo.

2.3.5 Testes de Estacionariedade

Se um processo possui média μ e a autocovariância γ que são independentes de t , então este processo é dito estacionário de covariância ou fracamente estacionário.

Existem diversos testes para avaliação da estacionariedade de uma série temporal. O teste KPSS (Kwiatkowski, Phillips, Schmidt, and Shin) é um teste reverso onde a hipótese nula é que a série é estacionária, assumindo-se assim um modelo similar ao da Equação 2.19 (HAMILTON, 1994).

2.3.6 Detecção e Caracterização de Oscilações

Em virtude das constantes demandas por redução de custos nos processos produtivos, diversas frentes de estudos foram abertas com o intuito de se identificar desperdícios e formas efetivas de eliminá-los. Com este foco, uma linha de pesquisa foi conduzida para detecção e caracterização de oscilações (*Oscillation, Detection and Characterization- ODC*)

em sinais presentes em malhas de controle, os quais podem ser representados por séries temporais com entradas exógenas ou não, visando eliminação imediata já que estas estão diretamente relacionadas com os custos de produção. As causas de oscilações em malhas de controle são as mais diversas como, por exemplo: má sintonia do controlador, equipamentos danificados ou até mesmo erro de projeto (DEPIZZOL, 2011).

Este método de caracterização e detecção de oscilação (ODC) foi desenvolvido por (KARRA; KARIM, 2009) e baseia-se no algoritmo apresentado por (THORNHILL; HUANG; ZHANG, 2003) e propõe melhorias utilizando a função de densidade espectral (*Power Spectral Density* - PSD) para identificar e separar as prováveis frequências de oscilação do sinal com foco em malhas de controle de processo industrial.

O método ODC é classificado como um método baseado na autocorrelação, uma vez que a avaliação do comportamento oscilatório do sinal é obtida a partir da função de autocorrelação (*Auto-Correlation Function* - ACF). O emprego conjugado de ferramentas matemáticas para identificação correta das oscilações na presença de adversidades (ex.: tendências, comportamentos transitórios, *outliers*, escolha do período de amostragem e múltiplas frequências de oscilação) deve ser considerado nos algoritmos de detecção.

O método ODC utiliza a PSD para identificar as frequências dominantes do sinal no tempo, a ACF para atenuar o ruído, cálculo dos períodos, e cálculos dos índices de detecção de oscilação e filtros que são utilizados para pré-tratamento do sinal e separação das suas múltiplas frequências, quando existentes. No método ODC, a análise do espectro é realizada a partir da PSD normalizada, observada na equação abaixo:

$$\Phi_{Nx}(f) = \frac{\Phi_x(f)}{\sum \Phi_x(f)} \quad (2.24)$$

As diferentes características do sinal no tempo determinarão a forma e distribuição das suas frequências em $\Phi_x(f)$. O número de frequências de oscilação, as amplitudes do sinal oscilatório, a intensidade do ruído, dentre outros fatores, determinarão a amplitude dos picos de potência em $\Phi_x(f)$. A amplitude de um sinal oscilatório presente em um sinal com múltiplas frequências pode ser encontrada através da soma da energia na banda de frequência correspondente em $\Phi_x(f)$, assim como mostrado na Equação (2.25).

$$P = \frac{\sum_{f=f_1}^{f_2} \Phi_x(f)}{\sum_{f=0}^{2/T} \Phi_x(f)} \quad (2.25)$$

onde f_1 e f_2 são os limites de frequência baixa e alta, respectivamente, selecionados em $\Phi_{Nx}(f)$. Quando a frequência analisada possui baixos valores de P , isto é indicativo que o sinal não possui atividade oscilatória significativa na banda selecionada, i.e., o comportamento oscilatório do sinal é dominante para outras frequências. Um limiar de potência denotado por ϵ_p foi estabelecido para classificar as bandas de frequência em $\Phi_{Nx}(f)$ relevantes na análise do comportamento oscilatório do sinal. As frequências inferiores (f_L) e superiores (f_U) das bandas selecionadas serão determinadas pela intersecção ϵ_p com $\Phi_{Nx}(f)$ (DEPIZZOL, 2011).

Os valores de ϵ_p podem ser selecionados visando identificação de sinais a partir da amplitude (energia) desejada. A Figura 7 mostra este corte para um valor de $\epsilon_p = 0, 1$.

Múltiplas frequências de oscilação introduzem cruzamentos por zero irregulares, dificultando a análise oscilatória do sinal. Sendo assim, é realizada a filtragem das frequências vizinhas aos picos de amplitude identificados como dominantes (filtros passa-faixa), separando-se o sinal em diversos pacotes para análise da ACF separadamente, conforme pode ser visto na Figura 8.

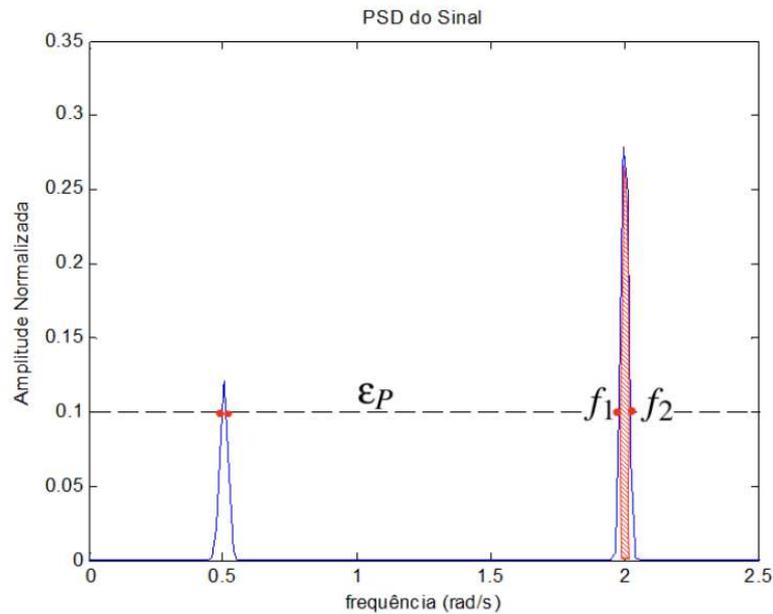


Figura 7: Seleção das bandas de frequências dominantes. (DEPIZZOL, 2011).

Mesmo com característica da PSD de distribuir sua potência de acordo com as frequências de oscilações dominantes no sinal, ela não é recomendada para determinar seus períodos de oscilações (THORNHILL; HUANG; ZHANG, 2003). A medição dos períodos de oscilação (T_p) do sinal pode ser obtida pelo dobro das distâncias entre os sucessivos cruzamentos por zeros do sinal autocorrelacionado.

O índice de regularidade de períodos (r) utiliza o conceito de desvio padrão para verificar o quão regular são os valores dos períodos calculados, servindo de indicativo para caracterizar a presença de comportamento oscilatório no sinal, sendo calculado por,

$$r = \frac{1}{3} \frac{\overline{T_p}}{\sigma_{T_p}} \quad (2.26)$$

onde $\overline{T_p}$ é a média dos períodos e σ_{T_p} o desvio padrão dos períodos. A obtenção de r utiliza somente os dez primeiros períodos de oscilação calculados, pois a magnitude da função de autocorrelação fica distorcida para atrasos grandes quando o sinal não é perfeitamente oscilatório. Uma oscilação é considerada regular se o desvio padrão dos períodos é menor que um terço do seu valor médio, ou seja, $r \geq 1$ (DEPIZZOL, 2011).

As oscilações tendem a tornar-se cada vez mais regulares ao passo que se propagam pelo processo, já que os equipamentos da planta industrial funcionam como filtro passa-baixas.

Desse modo, o algoritmo para detecção de oscilações é formado pelas seguintes etapas:

- Transformar o sinal em um sinal de média zero e normalizado;
- Aplicar filtro passa-banda com largura de $[0,02 - 0,99]$ Hz/Hz para remover as componentes de frequências muito altas (ex.: ruído de medição) e muito baixas;
- Gerar a PSD normalizada e identificar as frequências dominantes (picos) com valores superiores a uma amplitude escolhida da energia total do sinal (ϵ_p em percentual);

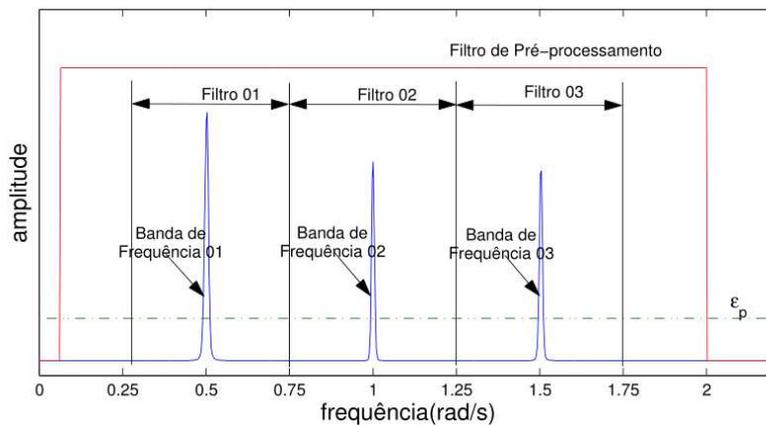


Figura 8: Representação do cálculo das bordas dos filtros passa-banda a partir da PSD, para um sinal com múltiplas frequências de oscilação (DEPIZZOL, 2011).

- No caso de detecção de múltiplas frequências dominantes, aplicar filtros passa-faixa englobando a vizinhança de cada frequência dominante para gerar diferentes conjuntos de dados;
- Gerar ACF e calcular a regularidade para cada pacote de oscilações identificadas para cada frequência dominante;
- As oscilações com energia acima do ponto de corte ϵ_p e que possuam $r > 1$ são consideradas oscilações significativas e são apresentadas estas frequências, seu valor de r e sua energia percentual em relação ao total do sinal.

2.4 Séries Temporais Usando Redes Neurais Artificiais

Modelagem de séries temporais não-lineares vem recebendo interesse crescente tanto no ponto de vista teórico quanto de aplicações. Apesar de existir uma gama grande de métodos disponíveis para auxiliar o pesquisador durante o processo de identificação do tipo de não-linearidade a ser modelado, infelizmente, não são tão raros casos de seleção de modelos não-lineares sem uma teoria científica sólida. Assim, a seleção de um modelo não-linear adequado ou seja, uma correta identificação do tipo de não-linearidade, é um problema real nas análises de séries temporais (LUUKKONEN; SAIKKONEN; TERASVIRTA, 1988)

As redes neurais artificiais (RNAs) são estruturadas de forma a modelar as capacidades de processamento dos sistemas nervosos. Os diversos estudos que vem surgindo nesta área somados às numerosas aplicações sugerem que as RNAs são fortes candidatos quando se trabalha com previsão de séries temporais. Em contramão aos métodos baseados em modelos tradicionais, as RNAs são guiadas por dados, autoadaptativas, não-lineares, usam métodos estatísticos não-paramétricos e há em seu conceito, poucas afirmativas a priori em relação aos modelos do problema em estudo. A forma pode ser de grande valia em processos não-lineares que possuem funções e características desconhecidas e que, como resultado, são difíceis de estimar (JHA; SINHA, 2014) e (ZHANG; QI, 2005).

2.4.1 Previsão de Séries Temporais

O desenvolvimento de redes neurais artificiais como ferramenta para a previsão de séries temporais não-lineares é bem usual (KHASHEI; BIJARI, 2010), (ZHANG; PATUWO; HU, 2001), (ZHANG; QI, 2005) e (HU; HWANG, 2002). A principal diferença em relação às técnicas tradicionais utilizadas para previsão é sua estrutura flexível e que elas podem aprender padrões lineares ou não-lineares em séries temporais e prevê-los com precisão em casos que as técnicas tradicionais falham. Sua natureza paralela protege o modelo de falhar totalmente mesmo se um elemento da rede neural falhar, o sistema continuará sua tarefa. Sua operação para previsão de séries temporais pode ser facilmente descrita como os dados da série temporal como entrada de uma caixa preta, contendo um processo conhecido pelo programador/pesquisador, entregando na saída as previsões desta série temporal (ver Figura 9). Uma variedade de técnicas de RNAs foram propostas, investigadas e bem-sucedidas para previsão de séries temporais (DONATE; LI, 2013) e (AGUILAR; TURIAS; M.J.JIMENEZ-COME, 2014), sendo algumas delas inclusive com modelos híbridos com séries temporais como em (LIU; TIAN; LI, 2012), (FARUK, 2010) e (KHASHEI; BIJARI, 2011).

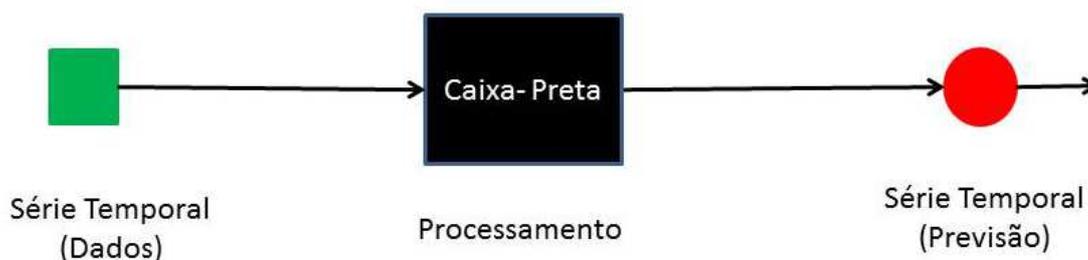


Figura 9: Processo de Previsão de Séries Temporais com RNAs.

2.4.2 Uma Abordagem Biológica

A definição tradicional do termo rede neural veio dos neurônios biológicos de que o cérebro humano é composto, que estão densamente interconectados. Muitas criaturas vivas, incluindo seres humanos, tem a habilidade de se ajustar ao ambiente que encontra-se em constante mudança. Eles conseguem isto através de sistemas de controle que possuem a habilidade de aprender. Este sistema nos seres humanos é o cérebro, que é formado por bilhões de neurônios interconectados.

Cada neurônio consiste de um corpo celular (soma), um axônio e os dendritos, conforme pode ser visualizado na Figura 10. Dendritos recebem sinais de entrada que são somados de forma ponderada, sendo então esta entrada processada através do corpo celular. O Foco dos axônios é a transmissão dos sinais eletroquímicos para outros neurônios. No final do axônio há terminações que convertem o sinal eletroquímico em informação química antes de passar para outros neurônios. Os neurônios são conectados através de estruturas chamadas sinapses e estas são cruciais para uma boa transmissão dos sinais químicos entre eles. Cada sinapse, dependendo do sinal químico transmitido, é habilitada para amplificar ou atenuar a ligação.

Na tentativa de explicar quão preciso e complicado é um sistema cerebral descrito acima, Haykin afirmou que o cérebro é um computador altamente complexo, que opera em paralelo e de forma não-linear(HAYKIN, 1998).

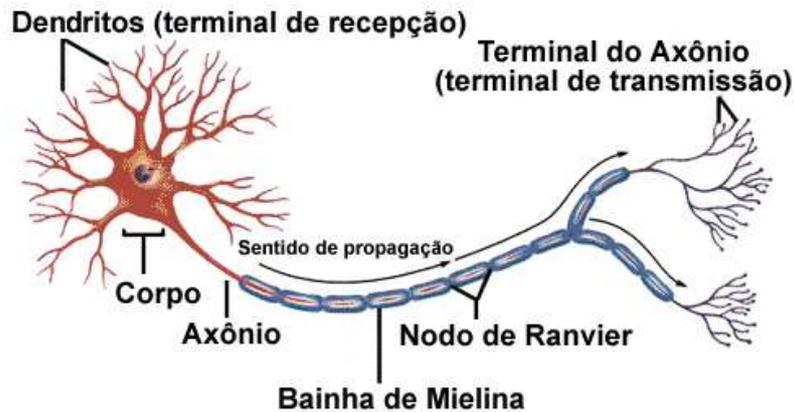


Figura 10: Neurônio Biológico.

O primeiro modelo de redes neurais artificiais que foi inspirado em estruturas de neurônios biológicos, foi apresentado pelo neurofisiologista americano Warren McCulloch e logicista americano Walter Pitts. O modelo deles era baseado na regra do ‘tudo ou nada’ que afirma que a resposta de um neurônio a um estímulo externo é independente da força do estímulo mas existe um gatilho de ativação que, somente quando é superado, ativa o neurônio. Isto sugere que a rede neural definida por eles é um sistema de entrada-saída. Alguns destes canais de entrada agem como estímulos para excitar os neurônios, enquanto outros restringem o processo. O gatilho é, então, estimulado pela soma das entradas de excitação (MCCULLOCH; PITTS, 1943). Assim, entradas são considerados estados binários com 1 representando canais de entrada que causam excitação no neurônio e 0 os canais sem excitação. Esta idéia se baseia nos operadores lógicos matemáticos Booleanos. Na Figura 11 pode-se verificar a configuração básica de um neurônio artificial McCulloch e Pitts - MCP.

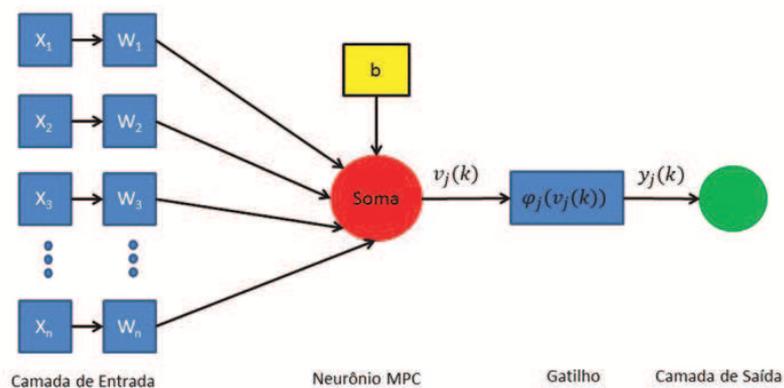


Figura 11: Neurônio Artificial MCP (McCulloch e Pitts).

Este modelo é equacionado da seguinte forma:

$$v_j(k) = \sum_{i=0}^n X_i w_i + b y_j(k) = \phi_j(v_j(k)), \quad (2.27)$$

onde n é o número de sinais de entrada do neurônio, X_i é o i -ésimo sinal de entrada do neurônio, w_i é o peso associado com o i -ésimo sinal de entrada, b é o limiar de cada neurônio, $v_j(k)$ é a resposta ponderada do j -ésimo neurônio em relação ao instante k ,

$\phi_j(\cdot)$ é a função de ativação para o j -ésimo neurônio e $y_j(k)$ é o sinal de saída do j -ésimo neurônio em relação ao instante k .

Cada neurônio artificial é capaz de computar os sinais de entrada e a respectiva saída. A função de ativação usada para calcular o sinal de saída é tipicamente não-linear (em geral a função sigmoide, tangente hiperbólica, logarítmica ou Gaussiana). O processo de ajuste dos pesos w_j associados ao j -ésimo neurônio de saída é feito pelo cálculo do sinal de erro em relação à k -ésima iteração ou ao k -ésimo vetor de entrada. Este sinal de erro é calculado pela seguinte equação:

$$e_j(k) = d_j(k) - y_j(k) \quad (2.28)$$

onde, $d_j(k)$ é a resposta desejada do j -ésimo neurônio de saída. Somando todos os erros quadráticos produzidos pelos neurônios de saída da rede em relação à k -ésima iteração, tem-se:

$$E(k) = \frac{1}{2} \sum_{i=0}^p e_i^2(k) \quad (2.29)$$

onde p é o número de neurônios da saída. Em busca de uma configuração de pesos ótima, $E(k)$ é minimizado pelo ajuste dos pesos sinápticos w_j .

2.4.3 Arquitetura

As diferentes arquiteturas de RNAs são formadas pela combinação de neurônios artificiais e são definidas pelo tipo de conexão entre os neurônios. Existem dois tipos baseados nestas conexões: as RNAs com alimentação para frente (*feedforward*) e recorrentes (*feedback*). Redes *feedforward*, como o nome já diz, possuem a característica que as camadas de entrada são direcionadas para a camada de saída, mas não o contrário. Estas são as formas mais simples de redes neurais e são divididas em dois tipos: camadas simples ou multi-camadas. As RNAs multi-camadas possuem camadas intermediárias chamadas de camadas ocultas que ficam entre as camadas de entrada e saída, sendo também conhecidas como RNAs multi-perceptron, que podem ser visualizadas na Figura 12(a).

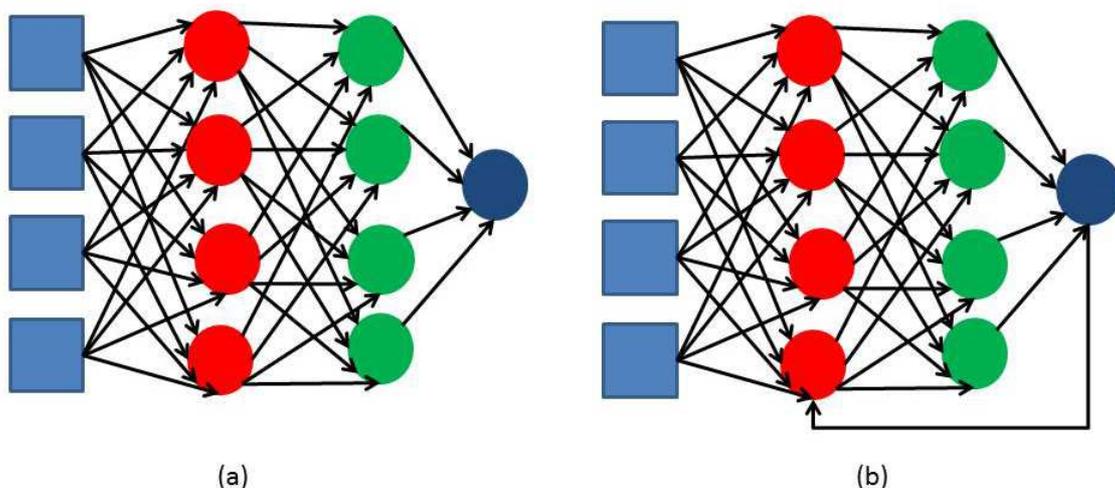


Figura 12: Redes Neurais Artificiais - Rede *Feedforward* (a) e Rede *Feedback* (b).

Nestas RNAs, cada neurônio transfere o seu sinal apenas para os neurônios que se encontram em uma das camadas subsequentes, ou seja, não há retro acoplamento entre os neurônios. Três tipos de camadas são identificadas:

- Camada de Entrada: é a interface de entrada, onde os sinais externos de entrada são passados para dentro da rede;
- Camadas Ocultas: definem a representação interna do mapeamento e não tem ligação direta com o sinal externo;
- Camada de Saída: os sinais de saída da rede são capturados a partir dos nós de saída.

As RNAs *Feedback* por sua vez, além das camadas de entrada, ocultas e saída, possuem componentes cíclicos, ou seja, as saídas realimentadas as entradas. Este tipo de rede pode ser visualizado na Figura 12 (b).

2.4.4 Treinamento

RNAs possuem uma relação bidirecional com o ambiente. Nas funções da rede neural, os pesos das sinapses entre os neurônios são inicialmente configurados aleatoriamente. Assim, há casos em que a performance da rede neural não é satisfatória, apresentando erros significativos. Nestes casos, é aconselhável ajustar os pesos, sendo este procedimento conhecido como treinamento. Há diversas formas de se treinar RNAs visando melhorar suas performances.

O treinamento supervisionado é o mais usado atualmente e pode ser realizado por um professor externo, o programador da rede neural, ou a rede sozinha (aprendizado por autosupervisão). No treinamento supervisionado, o programador busca por saídas onde ele tem as entradas mapeadas. Dentro desta categoria, a técnica de aprendizagem mais popular para treinamento de RNAs multi-camadas *feedforward* é a retro-propagação (*backpropagation*).

Na primeira fase do treinamento, o sinal de entrada é propagado para a frente, das entradas até a saída da rede. Como o valor da saída desejada $y_j(k)$ para a entrada corrente $x_j(k)$ é conhecido, o erro para a camada de saída pode ser calculado. Como não existem valores de saída desejados para as camadas intermediárias da rede, somente para a camada de saída, o ajuste do peso das camadas intermediárias é feito através da propagação para trás do erro da camada de saída, o que caracteriza o treinamento como *backpropagation* (REZENDE, 2003).

2.4.5 Redes Neurais com Atraso de Tempo

As redes neurais com atraso de tempo (*time-delay neural networks* - TDNN) são amplamente utilizadas para previsão de séries temporais devido ao seu tipo de processamento. Dado os valores observados de um sistema dinâmico no tempo menor ou igual a t , o problema de previsão consiste na utilização destes dados observados para prever $x(t+p)$, onde x representa o estado do sistema dinâmico e p é o horizonte de previsão. A Figura 13 ilustra como a sequência finita $\{x(t), x(t-1), x(t-2), \dots, x(t-n)\}$ é mapeada para uma saída única y . Nesta TDNN, cada conexão é configurada com interval específico no passado. A primeira conexão é configurada no instante atual, a segunda é configurada um instante no passado e assim por diante. As TDNN são funcionalmente equivalentes a filtros de resposta ao impulso finito (*Finite Impulse Response* - FIR) e são conhecidas redes recorrentes (*feedforward*) (HU; HWANG, 2002).

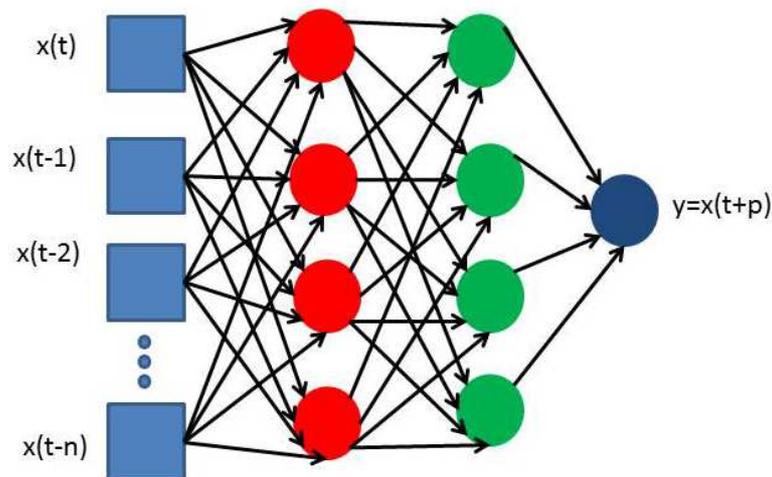


Figura 13: Arquitetura de uma Rede Neural com atraso de tempo.

2.4.6 Redes Neurais Recorrentes

As redes neurais recorrentes (*Recurrent Neural Networks* - RNNs) possuem conexões retroalimentadas. Elas capturam as relações entre as entradas atrasadas mantendo os estados que possuem memória. As RNNs são eficazes no aprendizado de dependências temporais em processos de curta memória. Em processos de memória longa, as RNNs são menos eficazes pois os erros tendem a se diluir ao passar nas camadas muitas vezes. Devido sua natureza dinâmica, RNNs foram bastante usadas em previsão de séries temporais. Treinar RNNs tende a ser difícil devido suas conexões realimentadas que causam tempo elevado de treinamento. (HU; HWANG, 2002). Na Figura 14, pode-se visualizar uma RNN típica.

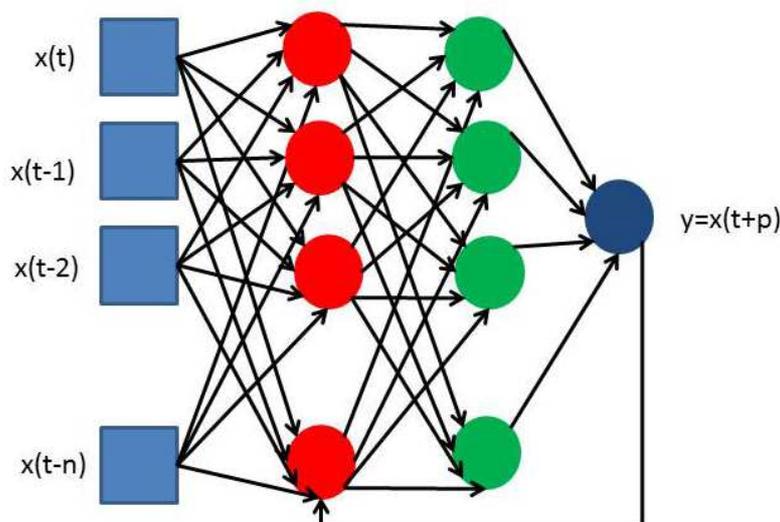


Figura 14: Arquitetura de uma Rede Neural com atraso de tempo e recorrentes.

2.5 Estimação de Parâmetros de Séries Temporais

A estimação de parâmetros está diretamente relacionada com a otimização pois, dependendo do algoritmo adotado, pode-se melhorar ou não o resultado das previsões. Os algoritmos numéricos para solução de problemas de otimização são classificados em métodos de programação matemática e métodos probabilísticos.

Os métodos de programação matemática são classificados em métodos de programação linear, programação não-linear e métodos baseados em teoria de aproximações como programação linear seqüencial e programação quadrática seqüencial (*Sequential Quadratic Programming* - SQP). Estes métodos também são classificados em métodos para solução de problemas de otimização sem restrição e com restrição.

Em problemas de otimização, existem diversos algoritmos de programação matemática, em adição aos algoritmos evolucionários, que são definidos de acordo com as características do problema (função-objetivo e restrições). Os algoritmos de programação matemática, restrita e irrestrita, são procedimentos iterativos em que novos pontos x são gerados a partir do ponto atual x_0 conforme equação abaixo:

$$x = x_0 + td \quad (2.30)$$

Pela Equação (2.30), nota-se os algoritmos podem ser divididos em duas etapas principais: a determinação da direção de busca d e a avaliação do parâmetro escalar t que representa o tamanho do passo a ser dado ao longo da direção de busca. A partir desta equação, vários algoritmos podem ser construídos utilizando diferentes técnicas para a determinação da direção de busca e do tamanho do passo (PEREIRA, 2002).

A busca por mínimos globais é difícil mesmo quando não há restrições. A situação pode ser minimizada quando há inclusão de restrições desde que estes excluam mínimos locais e facilitem a busca pelos mínimos globais na região da busca. Contudo, restrições podem também fazer a busca muito mais difícil.

2.5.1 Otimização sem restrições

Considere como exemplo o problema de minimização de uma função,

$$\text{minimizar} \quad f(x), \quad x \in R^n \quad (2.31)$$

A função $f(x)$ da Equação (2.31) a ser minimizada é chamada de função-objetivo.

Para que um ponto qualquer x_p seja um mínimo local do problema acima, temos como condição que o gradiente da função-objetivo em x seja nulo ($\nabla f(x_p) = 0$) e que a matriz Hessiana $\nabla^2 f(x_p)$ seja positiva, ou seja:

$$\langle d, \nabla^2 f(x_p) d \rangle > 0 \quad \forall d > 0 \quad (2.32)$$

Se $f(x)$ é uma função convexa definida em R^n e ω é um conjunto de pontos $x \in R^n$ onde $f(x)$ está sobre o mínimo. Então ω é convexo e todo mínimo local é também mínimo global.

Se $f(x)$ é uma função convexa e existe $x_p \in R^n$ tal que para todo $y \in R^n$,

$$\langle \nabla^2 f(x_p), (y - x_p) \rangle > 0 \quad (2.33)$$

então x_p é um ponto de mínimo global desta função $f(x)$ (SILVA, 1997).

Entre os métodos de otimização sem restrições encontra-se o método de Newton. Neste método, a função custo $f(x)$ é expandida até a segunda ordem pela série de Taylor em torno do ponto x_0 :

$$f(x) = f(x_0) + \nabla f(x_0)(x - x_0) + \frac{1}{2}(x - x_0)^t \nabla^2 f(x_0)(x - x_0) \quad (2.34)$$

se

$$d = \delta x = (x - x_0) \longrightarrow x = d + x_0 \quad (2.35)$$

e

$$g = \nabla f(x_0) \quad e \quad H = \nabla^2 f(x_0) \quad (2.36)$$

Substituindo-se (2.32) e (2.33) em (2.31), temos:

$$f(d + x_0) = f(x_0) + d^t g + \frac{1}{2} d^t H d \quad (2.37)$$

onde d é o incremento de x_0 , g é vetor gradiente de f e H é uma matriz simétrica positiva definida como a hessiana da função f no ponto x_0 . Pode-se notar que a Equação (2.34) é uma função quadrática sobre a variável d . Este algoritmo busca reduzir a função-custo de tal forma que $f(d + x_0)$ seja menor que $f(x_0)$ a cada passo minimizado através da equação quadrática (PEREIRA, 2002).

Contudo, o esforço computacional para fatorização das matrizes Hessianas pode ser proibitivo em casos onde a quantidade de dados a ser processada é muito grande. Nestes casos utilizam-se métodos denominados Quase-Newton que resolvem o problema sem perder as propriedades de convergência do método de Newton. Nestes métodos, as Hessianas são aproximadas pelos gradientes durante as iterações. Um dos métodos de destaque nesta categoria é o método BFGS (Broyden - Fletcher - Goldfarb - Shanno) (NOCEDAL; WRIGHT, 2006).

2.5.2 Otimização com restrições

Considere agora como exemplo o problema de minimização de uma função desta forma,

$$\begin{aligned} &\text{minimizar} && f(x), \quad x \in R^n \\ &\text{sujeito a} && g_i(x) = 0, \quad i \in E \\ &&& g_i(x) \leq 0, \quad i \in I \end{aligned} \quad (2.38)$$

onde $f(x)$ é chamada de função-objetivo e $g_i(x)$ são restrições. O conjunto E é chamado de conjunto de restrições de igualdade, enquanto que o I é chamado de conjunto de restrições de desigualdade.

Se x_p é um ponto que satisfaz as condições $g_i(x)$, $i \in E \cup I$ e seus gradientes ($\nabla g + i(x_p)$, $i \in E \cup I$) são linearmente independentes este ponto é dito ponto regular.

A função Lagrangeana é definida por:

$$L(x, u) = f(x) + \sum_{i=1}^m u_i g_i(x), \quad (2.39)$$

onde $u_i \in R$, $i = 1, \dots, m$ são os multiplicadores de Lagrange.

Já o vetor gradiente de $L(x, u)$ é representado por $\nabla L(x, u) = \begin{bmatrix} \nabla_x L(x, u) \\ \nabla_u L(x, u) \end{bmatrix}$, onde ∇_x e ∇_u indicam as derivadas parciais em relação a x e a u , respectivamente.

O vetor gradiente da função Lagrangeana em relação à x é dado por:

$$\nabla_x L(x_p, u_p) = \nabla f(x_p) + \sum_{i=1}^m u_{ip} \nabla g_i(x_p), \quad (2.40)$$

onde u_p é o vetor dos multiplicadores de Lagrange no ponto ótimo.

Assim, pode-se obter a matriz Hessiana através da função Lagrangeana para este mesmo ponto:

$$\nabla_x^2 L(x_p, u_p) = \nabla^2 f(x_p) + \sum_{i=1}^m u_{ip} \nabla^2 g_i(x_p). \quad (2.41)$$

As condições que Kuhn-Tucker estabelecem que se um ponto mínimo local x_p da Equação (2.32) é regular, então existem multiplicadores de Lagrange u_p de tal forma que as seguintes equações são satisfeitas:

$$\begin{aligned} \nabla f(x_p) + \sum_{i=1}^m u_{ip} \nabla g_i(x_p) &= 0 \\ g_i(x_p) &= 0, \quad i \in E \\ g_i(x_p) &\leq 0, \quad i \in I \\ u_{ip} &\leq 0, \quad i \in I \\ u_{ip} g_i(x_p) &= 0, \quad \forall i \end{aligned} \quad (2.42)$$

Além das condições de Kuhn-Tucker (SILVA, 1997), para que um ponto x_p seja um mínimo local, é necessário que seja um ponto regular e que sua Hessiana seja positiva, ou seja:

$$\langle d^T, \nabla_x^2 L(x_p, u_p) d \rangle \geq 0 \quad \forall d \in G_p \quad (2.43)$$

onde G_p é:

$$\begin{aligned} \{d/d \neq 0, d^T \nabla_{g_i}(x_p) = 0, i \in (E \cup I)/u_{ip} > 0 \\ \text{e} \quad d^T \nabla_{g_i}(x_p) \leq 0, i \in I/u_{ip} = 0\} \end{aligned}$$

No caso particular em que temos uma função-objetivo e das restrições serem convexas (2a ordem), e as funções $f(x)$ e $g_i(x)$ são contínuas com derivadas parciais contínuas de primeira ordem e as condições de Kuhn-Tucker são atendidas para x_p , então o ponto x_p é uma solução global do problema. Neste caso, a programação é dita programação convexa (SILVA, 1997).

2.5.3 Programação Sequencial Quadrática

A programação sequencial quadrática (*Sequential Quadratic Programming* - SQP) é um dos mais bem-sucedidos métodos para solução numérica de problemas de otimização não-linear (*Non-Linear optimization Problems* - NLP) com restrições, inclusive, é o método utilizado pelo software matemático MATLAB para identificação de parâmetros de séries temporais. Nestes tipos de problemas não-lineares, como mostrado na Equação

(2.38), estão contidos os casos especiais que são os problemas de programação linear e quadrática, onde f é linear ou quadrático e as funções de restrições são afins.

O SQP é um procedimento iterativo cujos modelos de NLP são dados através da solução de um subproblema de programação quadrática (*Quadratic Programming* - QP) a cada iteração x^k , $k \in N_0$, usando esta solução para construir uma nova iteração x^{k+1} . Este processo é guiado de forma que a sequência $(x^k)_{k \in N_0}$ converge para um mínimo x_p do problema de otimização não-linear quando k tende a infinito. Os métodos de Newton e Quasi-Newton são utilizados para a solução numérica do sistema de equações algébrico não-linear. Contudo, a presença de restrições traz dificultantes de análise e implementação dos métodos SQP.

Um fato importante a ser considerado quanto ao uso da SQP é que se a matriz Hessiana G da Equação (2.38) é definida positiva, pode-se afirmar que trata-se de uma programação quadrática convexa, trazendo uma dificuldade similar à programação linear. No caso de programações quadráticas não-convexas, em que G é uma matriz indefinida, o trabalho torna-se mais desafiador pois pode haver alguns pontos estacionários ou mínimos locais (NOCEDAL; WRIGHT, 2006)

2.5.4 Algoritmos Genéticos

Os problemas de otimização do mundo real são complexos já que, na maioria dos casos, não apresentam boas propriedades matemáticas como continuidade, convergência, diferenciabilidade e modo único de operação. Contudo, estas propriedades são necessárias para utilização de algoritmos de otimização. Como resultado, a resolução de problemas gerais de otimização se tornou um tópico de estudo desafiador dentro do domínio das ciências computacionais e de otimização. Algoritmos evolucionários ou computação evolutiva apresenta um longo histórico de sucesso na resolução de problemas de otimização independentemente de terem ou não boas propriedades matemáticas (ELSAYED; SARKER; ESSAM, 2013).

Estes algoritmos evolucionários, dos quais os algoritmos genéticos são subclasses, cujas pesquisas tiveram início na década de 1950, trata de sistemas para a resolução de problemas que utilizam modelos computacionais baseados na teoria da evolução natural (seleção e hereditariedade). Eles partem do pressuposto que, numa dada população, os indivíduos com boas características genéticas tem maiores chances de sobrevivência e de produzirem indivíduos cada vez mais aptos. Como resultado, os indivíduos menos aptos tenderão a desaparecer. Assim, algoritmos genéticos (AGs) favorecem a combinação dos indivíduos mais aptos, ou seja, os candidatos mais promissores para a solução de um dado problema.

Quando os algoritmos genéticos são utilizados na resolução de problemas matemáticos, cada indivíduo da população corresponde a uma possível solução para o problema. Um mecanismo de reprodução, baseado em processo evolutivo, é aplicado sobre a população atual com o objetivo de explorar o espaço de busca e encontrar as melhores soluções (REZENDE, 2003).

Alguns exemplos de sucesso de utilização de algoritmos genéticos para estimação de parâmetros de modelos podem ser visto em (ABO-HAMMOUR et al., 2012), (FARAHAHAT; TALAAT, 2012), (FARAHAHAT; MCDONALD; NELSON, 1999), e (HAMEDANI; SAMET, 2010).

2.5.4.1 Seleção

A seleção é o componente que guia o algoritmo para a solução através de indivíduos escolhidos pela ordenação entre os mais aptos e os menos aptos. Isto pode ser uma operação determinística ou, em muitos casos, pode haver componentes aleatórios. Uma variação que é muito popular atualmente é a probabilidade de escolha de certo indivíduo ser proporcional à sua aptidão. Isto pode ser obtido através do cálculo abaixo:

$$P[b_{j,t} \text{ é selecionado}] = \frac{f(b_{j,t})}{\sum_{k=1}^m f(b_{k,t})} \quad (2.44)$$

onde $f(\cdot)$ é a função-custo do problema. Este sistema é, em sua essência, uma roleta onde as chances de escolha não são divididas igualmente, dependendo das probabilidades de cada uma. A Figura 15 representa graficamente este conceito.

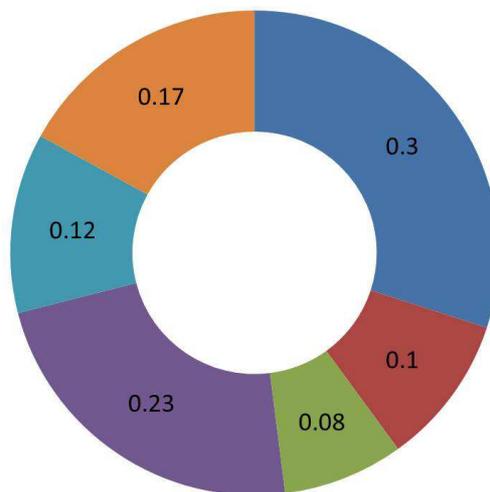


Figura 15: Representação gráfica da roleta de seleção, onde o número de alternativas m é 6. O número dentro dos arcos corresponde à probabilidade de cada alternativa ser selecionada.

2.5.4.2 Cruzamento

Em reprodução sexual, como acontece no mundo real, o material genético dos pais é misturado quando os gametas dos pais se encontram. Usualmente, cromossomos são aleatoriamente separados e juntos, tendo como consequência que alguns genes de uma criança vem do pai enquanto outros vem da mãe.

Este mecanismo é chamado cruzamento. É uma poderosa ferramenta para introduzir novo material genético e manter a diversidade genética, mas com a incrível capacidade de que bons pais podem gerar filhos ainda melhores. Algumas investigações concluíram que o cruzamento é a razão pela qual espécies que se reproduzem sexualmente se adaptam mais rapidamente que espécies assexuadas.

Basicamente, cruzamento é a troca de genes entre cromossomos de dois pais. Num caso mais simples, podemos realizar este processo cortando dois pedaços em uma posição aleatória e trocar os dois pedaços finais. Este processo é chamado de cruzamento de um ponto e pode ser visualizado na Figura 16.

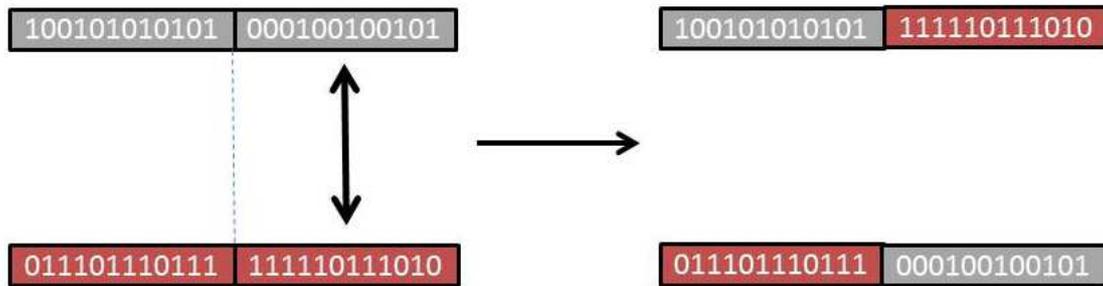


Figura 16: Exemplo de cruzamento binário de um ponto.

2.5.4.3 Mutação

O último ingrediente de um algoritmo genético simples é a mutação, a deformação aleatória de uma informação genética de um indivíduo através de radiação ou outras influências do meio-ambiente. Em uma reprodução real, a probabilidade de um gene sofrer mutação é quase a mesma para todos os genes. Sendo assim, é satisfatório considerar P_M a probabilidade de mutação de um gene ser modificado. É importante ressaltar que esta probabilidade de mutação deve ser pequena para evitarmos que o algoritmo genético possua comportamento caótico como uma busca aleatória.

A escolha do tipo de mutação depende do código e do problema analisados. Seguem algumas alternativas:

- Inversão de bit único: Com probabilidade P_M , um bit escolhido aleatoriamente é invertido;
- Inversão completa: Todos os dados são invertidos, bit a bit, com probabilidade P_M ;
- Seleção aleatória: Com probabilidade P_M , o dado é substituído por outro escolhido aleatoriamente.

2.5.4.4 Evolução Diferencial

A evolução diferencial (*Differential Evolution* - DE) é um algoritmo evolutivo desenvolvido por Rainer Storn e Kenneth Price em 1995, na tentativa de resolver o problema de ajuste polinomial de Chebyshev (STORN; PRICE, 1995). Ele foi desenvolvido como otimizador de parâmetros e valores de funções reais, podendo ser utilizado nos campos da engenharia, estatística e finanças, onde os problemas podem apresentar funções-objetivo não-diferenciáveis, não-contínuas, não-lineares, com ruído, multi-dimensionais, com mínimos locais, com restrições ou estocásticas (PRICE, 1999).

O algoritmo DE se inicia com uma população de NP candidatos a solução que pode ser representado como $X_{i,G} = 1, \dots, NP$, onde o índice i e G representam o indivíduo e a geração que aquela população pertence, respectivamente. O trabalho do algoritmo DE depende da manipulação e eficiência de seus três operadores: mutação, reprodução e seleção.

Mutação: o operador Mutação é o principal responsável pela diferenciação do algoritmo DE de outros algoritmos evolucionários. A operação de mutação no algoritmo DE aplica um vetor diferencial entre os membros da população existente para determinar o grau e a direção da perturbação aplicada ao indivíduo que irá sofrer a mutação. O processo

de mutação de cada geração começa selecionando aleatoriamente 05 indivíduos na população. As estratégias mais utilizadas de mutação dos códigos DE estão listadas abaixo:

1. DE/rand/1 $V_{i,g} = X_{r_1,g} + F \cdot (X_{r_2,g} - X_{r_3,g})$;
2. DE/rand/2 $V_{i,g} = X_{r_1,g} + F \cdot (X_{r_2,g} - X_{r_3,g}) + F \cdot (X_{r_4,g} - X_{r_5,g})$;
3. DE/best/1 $V_{i,g} = X_{best,g} + F \cdot (X_{r_1,g} - X_{r_2,g})$;
4. DE/best/2 $V_{i,g} = X_{best,g} + F \cdot (X_{r_1,g} - X_{r_2,g}) + F \cdot (X_{r_4,g} - X_{r_5,g})$;
5. DE/rand-to-best/1 $V_{i,g} = X_{r_1,g} + F \cdot (X_{best,g} - X_{r_2,g}) + F \cdot (X_{r_4,g} - X_{r_5,g})$;

onde, $i = 1, \dots, NP$ e $r_1, r_2, r_3, r_4, r_5 \in \{1, \dots, NP\}$ são aleatoriamente selecionados e satisfazem: $r_1 \neq r_2 \neq r_3 \neq i$ e $F \in [0, 1]$, F é um parâmetro de controle proposto por (STORN; PRICE, 1995).

Cruzamento: após a fase de mutação, o processo de cruzamento é ativado. A perturbação individual, $V_{i,G+1} = (v_{1,i,G+1}, \dots, v_{n,i,G+1})$, e o membro atual da população, $X_{i,G} = (x_{1,i,G}, \dots, x_{n,i,G})$, são submetidos à operação de cruzamento que gerará a população de candidatos, ou vetores de teste, $U_{i,G+1} = (u_{1,i,G+1}, \dots, u_{n,i,G+1})$, como segue:

$$u_{j,i,G+1} = \begin{cases} v_{j,i,G+1} & \text{se } rand_j \leq C_r \vee j = k \\ x_{j,i,G} & \text{caso contrário} \end{cases}$$

onde, $j = 1, \dots, n$, $k \in \{1, \dots, n\}$ é um índice aleatório por parâmetro e é escolhido uma vez para cada i . A taxa de cruzamento, $C_r \in [0, 1]$, outro parâmetro de controle do algoritmo DE que é escolhido pelo usuário.

Seleção: o processo de seleção do algoritmo DE também difere de outros algoritmos evolucionários. A população para a próxima geração é escolhida a partir de indivíduos da população atual e seu vetor de teste correspondente é selecionado através da seguinte regra:

$$X_{i,G+1} = \begin{cases} U_{i,G+1} & \text{se } f(U_{i,G+1}) \leq f(X_{i,G}) \\ X_{i,G} & \text{caso contrário} \end{cases}$$

Assim, cada indivíduo da população temporária (vetores-teste) é comparado com sua contra-parte da população atual. Aquele com menor valor para a função objetivo irá sobreviver ao torneio de seleção para a próxima geração. Como resultado, todos os indivíduos da próxima geração são tão bons ou melhores que suas contra-partes da geração atual. Os vetores-teste não são comparados contra todos os indivíduos da geração atual, mas somente contra um indivíduo, sua contra-parte, na geração atual.

O pseudo-código do algoritmo genético é mostrado a seguir:

- Passo 1 Realizar uma inicialização aleatória da população de pais;
- Passo 2 Calcular o valor da função-objetivo $f(X_i)$ para todo X_i ;
- Passo 3 Selecionar três indivíduos da população e gerar uma perturbação individual V_i usando uma das estratégias de mutação apresentadas anteriormente;
- Passo 4 Recombinar cada vetor x_i com a perturbação individual gerada no passo 3 para gerar um vetor-teste U_i usando a equação de cruzamento;

- Passo 5 Verificar se cada variável do vetor-teste está dentro dos limites. Se sim, ir para passo 6, se não, aplicar a equação $u_{i,j} = 2x_{min,j} - u_{i,j}$, se $u_{i,j} < x_{min,j}$ e $u_{i,j} = 2x_{max,j} - u_{i,j}$, se $u_{i,j} > x_{max,j}$;
- Passo 6 Calcular o valor da função-objetivo para o vetor U_i ;
- Passo 7 Escolher o melhor entre os dois (valor da função do alvo e vetor-teste) usando a equação de seleção para a próxima geração;
- Passo 8 Verificar, através dos critérios de convergência, se a meta foi alcançada e parar ou retornar ao passo 3;

Este algoritmo foi retirado do site da Universidade de Berkeley (STORN/CODE.HTML, 2015) e foi disponibilizado, além de outras linguagens, em Matlab. Foram realizadas algumas alterações no original visando adaptação ao problema específico.

3 NÍVEL DE LÍQUIDOS EM CADINHO DE ALTOS-FORNOS

3.1 Produção de Ferro-Gusa em Alto-Forno

O alto-forno pode ser considerado um trocador de calor e de massa em contra-corrente. O gás sobe pelo forno, enquanto a carga e o coque descem. O gás transfere calor para a carga e simultaneamente o oxigênio é transferido para o gás. Esta característica de processo em contra-corrente é bastante eficiente. O processo pode ser descrito da seguinte forma (GEERDES et al., 2007):

- Um alto-forno é carregado em camadas alternadas de coque e materiais que contêm ferro, conforme pode ser visto na Figura 17(WRIGHT et al., 2003);
- Insufla-se ar quente dentro do alto-forno por meio de ventaneiras. Uma ventaneira é um equipamento de cobre refrigerado por água que canaliza o ar quente (em torno de $1200^{\circ}C$) no interior do alto-forno na região denominada raceway;
- O ar quente gaseifica os combustíveis e comburentes (coque, ar, carvão, vapor, oxigênio, entre outros) através das ventaneiras. Neste processo, o oxigênio do ar se transforma em monóxido de carbono (CO). Este gás resultante tem alta temperatura de chama (acima de $2000^{\circ}C$) que consome o coque em frente às ventaneiras criando um vazio;
- O gás quente ascende por dentro do alto-forno, realizando um grande número de funções:
 1. Aquecimento do coque na área do ventre;
 2. Fusão da carga de minério de ferro criando vazios;
 3. Aquecimento do material na cuba do alto-forno;
 4. Eliminação de parte do oxigênio da carga através de reações químicas nas zonas de coesão;
 5. Ao fundir-se, o minério de ferro gera gusa e escória, que gotejam na zona de coque até o cadinho, que é retirado através de orifícios na carcaça do alto-forno, denominado furo-de-gusa.

O alto-forno possui uma forma típica de segmentos tronco-cônicos e cilíndricos. As seções do topo para baixo são: a goela, onde se encontra a superfície da carga; a cuba; o ventre (rampa paralela); a rampa e o cadinho. Uma visão geral dos principais equipamentos que formam o processo de um alto-forno pode ser visualizada na Figura 18 (WWW.IMAGES.ENCARTA.MSN.COM, 2010). Estes equipamentos são descritos a seguir:

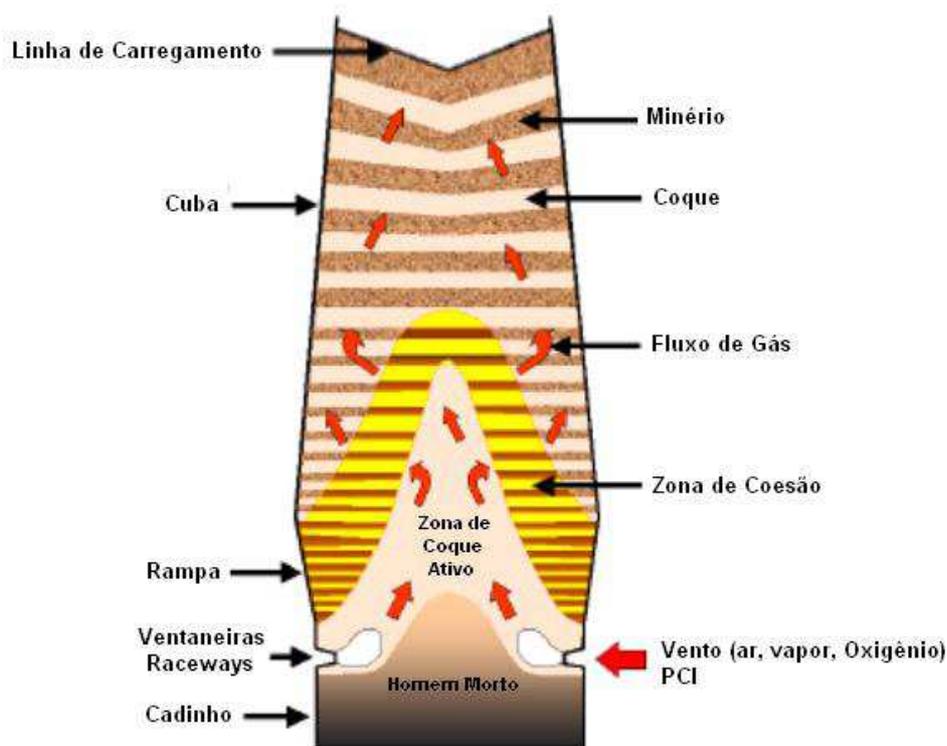


Figura 17: O processo do Alto-Forno.

Carregamento Região onde se encontram os silos de abastecimento do forno. Nestes silos estão armazenados coque, minério de ferro, pelotas e fundentes. Estes são transportados para o topo do forno pelas correias transportadoras ou caçambas de carregamento. É neste local onde são selecionadas as receitas (relações coque/minério) das cargas do alto-forno.

Topo do Forno É este o local por onde o alto-forno é carregado. Como o alto-forno opera com pressão positiva em relação à pressão atmosférica (entre 1, 5 e 3, 0kgf/cm²), o topo possui um sistema de pressurização /despressurização. Existem atualmente equipamentos com sistemas de duplo-cone e sistemas que operam com distribuidores rotativos (sem cone).

Limpeza de Gás Equipamento responsável pela coleta dos gases gerados no processo assim como a retirada das partículas sólidas contidas nos gases (cinzas de coque) através de sistemas de lavagem e decantação. É nesta região que é realizado o controle de pressão do topo do alto-forno.

Casa-de-Corrída Região do alto-forno onde é realizada a extração dos produtos finais do processo (gusa e escória) através de perfurações periódicas do cadinho do forno. Estes materiais são vazados em canais e separados por diferença de densidade. O gusa é então carregado em carros-torpedo ou carros-panela e transportado para a aciaria onde será transformado em aço.

Granulação de Escória A escória extraída do cadinho, na maioria dos casos, é granulada, após resfriada com água. A escória granulada pode ser utilizada para fabricação de cimento ou asfalto rodoviário.

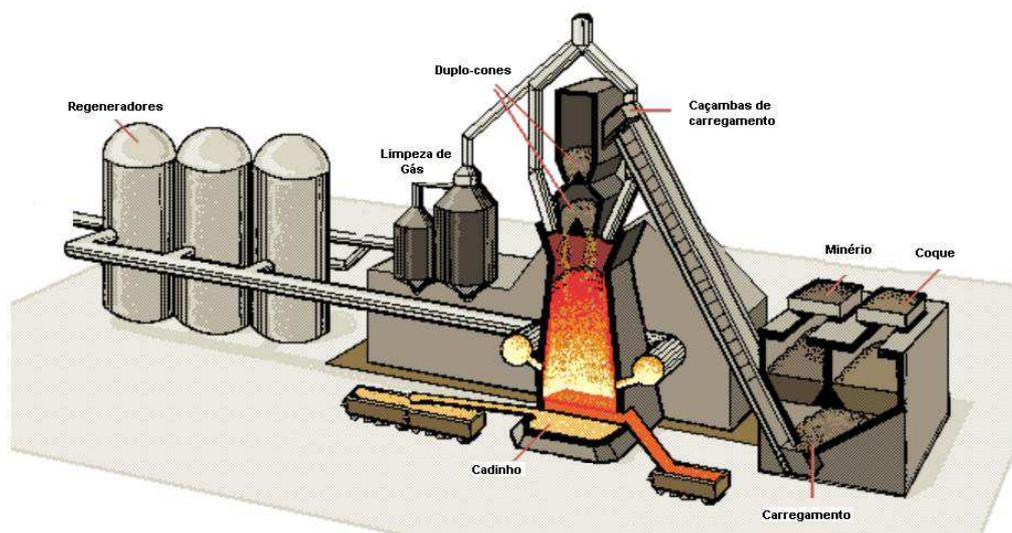


Figura 18: Visão geral de um alto-forno.

Regeneradores São trocadores de calor responsáveis por realizar um pré-aquecimento do ar que será injetado no alto-forno. Este ar quente (temperatura entre 1000 e 1300°C) é soprado através dos regeneradores para o forno. Neste equipamento é realizado, também, o controle da umidade e quantidade de oxigênio do ar. O transporte do ar quente é realizado pela linha de ar quente, anel de vento e ventaneiras.

Injeção de Carvão Pulverizado O uso de combustíveis auxiliares, que são injetados pelas ventaneiras, pode reduzir o custo do gusa. Os combustíveis auxiliares quase sempre são o carvão e o gás natural, porém o alcatrão e outros materiais podem também ser usados. No início dos anos 80 a injeção de óleo era mais comum. Isto mudou devido alteração dos preços relativos do carvão e óleo. O carvão é injetado por meio de lanças para dentro dos algaravizes, entra em ignição e é gaseificado. O equipamento de injeção de carvão pulverizado é formado pelos sistemas de armazenamento, moagem e transporte do carvão até o alto-forno.

Na Figura 19 podemos visualizar o rendimento médio de um alto-forno. Nota-se que para cada tonelada de ferro-gusa produzida, são necessários $1,6$ toneladas de minério-de-ferro, $0,3$ toneladas de coque, $0,2$ toneladas de oxigênio e 1000m^3 de ar. Nesta proporção de matérias-primas, são gerados 1150m^3 de gás de alto-forno e $0,3$ toneladas de escória, além do ferro-gusa.

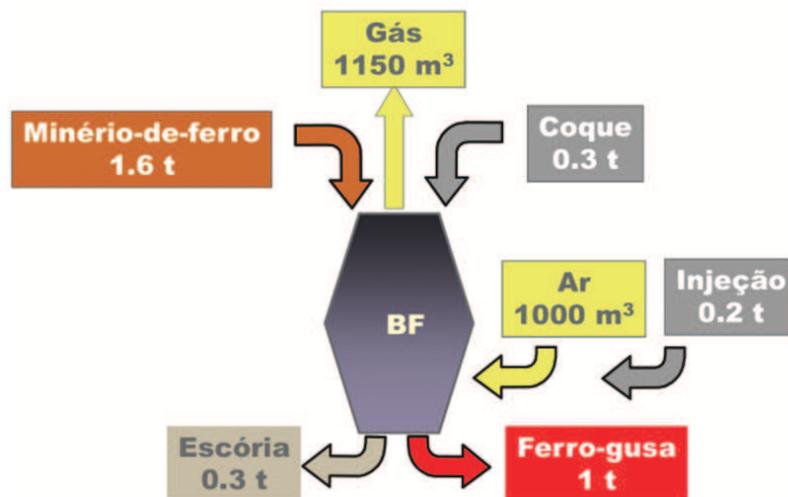


Figura 19: Redimento de um Alto-forno.

A carga do alto-forno desloca-se de cima para baixo numa velocidade, em condições normais de operação, constante. Para que a carga desça, vazios são criados durante a queima do coque na frente das ventaneiras e na fusão dos materiais no interior do forno. Os fenômenos de engaiolamentos (quando não há descida de carga) e arriamentos (descida rápida e descontrolada da carga) ocorrem quando há deficiências no processo de redução. As razões podem ser:

1. A força ascendente está muito alta;
2. Temperatura de operação muito alta;
3. Alterações na operação da casa-de-corrída (podendo afetar nível do líquido e flutuação da camada de coque no interior do cadinho);
4. Contenção da força descendente pela fusão dos materiais em zonas indevidas.

A resistência à vazão de gás num alto-forno carregado está localizada nas camadas de minério, considerando que a permeabilidade do minério é 4 a 5 vezes inferior à permeabilidade das camadas de coque. A velocidade de subida do gás é bem baixa (entre 2 e 5m/s) mas pode atingir velocidades próximas dos 10m/s no centro do alto-forno devido à maior concentração de coque nesta região. Podem existir situações que a fluidização do coque, e consequente elevação de velocidade de subida do gás, atinja a parte inferior do alto-forno ocasionando efeito chaminé (curto-circuito entre o topo e a parte baixa do forno). Neste caso, o gás escapa com temperatura muito alta para o topo e com baixa utilização de seu potencial redutor (GEERDES et al., 2007).

3.2 O Cadinho do Alto-Forno

O gusa e a escória líquidos são recolhidos no cadinho do alto-forno, bem abaixo das ventaneiras. O gusa e a escória não se misturam: a escória possui uma densidade mais baixa ($2.3t/m^3$) do que o gusa ($7.2t/m^3$) e flutua sobre ele. Na Figura 20 (GEERDES et al., 2007) pode ser visualizada uma representação esquemática de um cadinho e o furo-de-gusa. O furo-de-gusa apresenta pelo lado de dentro um cogumelo refratário com massa de fechamento do furo-de-gusa solidificada.

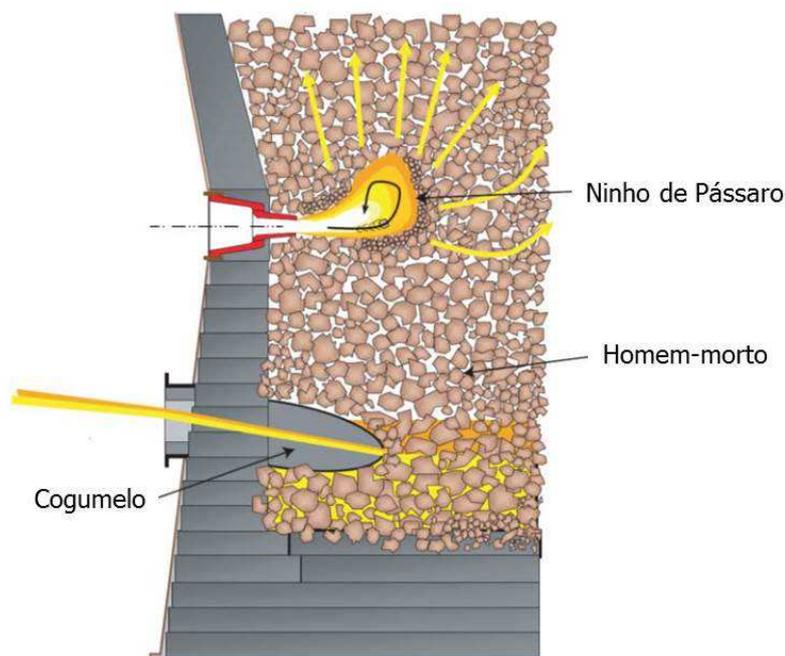


Figura 20: Cadinho de um alto-forno.

O alto-forno é vazado de 8 a 14 vezes por dia através do furo-de-gusa, o que corresponde a uma duração média de vazamento entre 1,5 e 3 horas. O tempo de residência da carga no interior do forno é de 5 a 6 horas. Assim, para um vazamento de 2 horas, aproximadamente $1/3$ da carga do forno é extraída. O início de sopro de gás no furo-de-gusa indica cadinho vazio, sendo necessário o fechamento imediato pois o ar soprado pelas ventaneiras se conectou com o furo-de-gusa.

Os níveis dos líquidos no cadinho afetam o processo de produção do alto-forno de duas formas:

- Os líquidos no cadinho afetam a descida de carga: quanto mais alto for o nível do líquido, mais fortemente o coque submerso empurrará a carga em sentido ascendente.
- Se a escória alcançar o nível das ventaneiras e não puder ser drenada, o fluxo de gás será severamente afetado. Isto pode causar desde um resfriamento do forno até uma obstrução total das ventaneiras demandando uma parada na planta para reparos.

Sendo assim, o nível dos líquidos no cadinho tem que ser mantido sob controle e, se possível, constante e baixo. Mesmo com uma boa prática de casa de corrida existe uma variação no nível líquido do cadinho de até 1 metro no interior do alto-forno.

Um alto-forno moderno possui pelo menos 2 (dois) furos-de-gusa. Os fornos grandes com cadinhos maiores que 10 metros de diâmetro, possuem geralmente 4 furos-de-gusa para melhorar o esgotamento dos líquidos de seu interior. O gusa é vazado no canal principal enquanto a escória escoá pelos canais secundários. O gusa e a escória são separados facilmente pela diferença de densidade. A maioria dos altos-fornos de alta produtividade são vazados de forma alternada e simétrica, ou seja, quando um furo-de-gusa é fechado,

o outro simetricamente oposto é aberto. O intervalo de tempo entre vazamentos pode ser reduzido a zero ou ocorrer aberturas paralelas em virtude de desvios operacionais.

A operação ideal da casa de corrida para um alto-forno grande e de alta produtividade é o vazamento contínuo, de forma alternada, com tempos de vazamentos semelhantes para as corridas e um fluxo de escória quase contínuo. As vazões de gusa e escória de um furo-de-gusa dependem das características físico-químicas do gusa e escória, do comprimento e diâmetro do furo-de-gusa e da pressão interna do forno. Uma vez que o furo se desgasta continuamente, especialmente pelo ataque da escória, as vazões de gusa e escória não são constantes. Além disso, no início da corrida apenas o gusa é extraído. Um padrão típico de fluxo de vazamento é visto na Figura 21 (GEERDES et al., 2007).

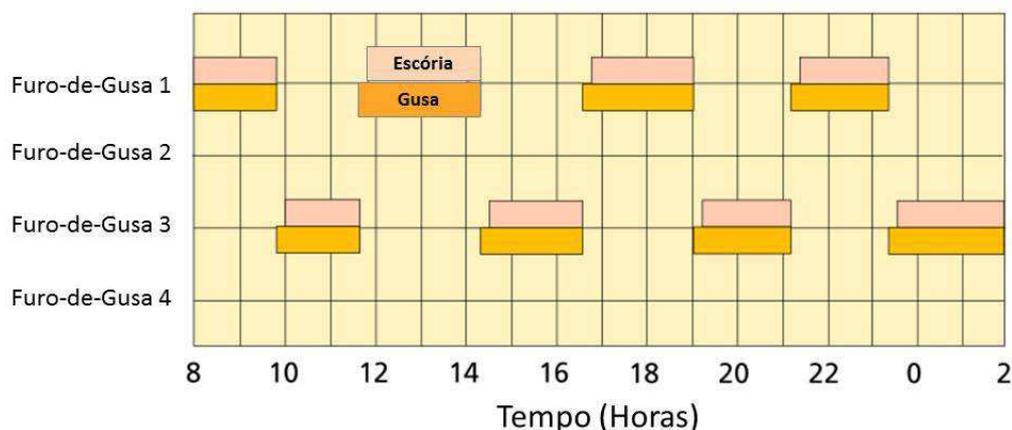


Figura 21: Padrão típico de operação de casa de corrida de um alto-forno.

O esgotamento do ferro-gusa do cadinho do alto-forno é um procedimento inerente à atividade produtiva, mas que diminui a vida útil do refratário nesta região, uma vez que envolve impactos e vibrações elevados nas paredes do alto-forno. Atualmente, as aberturas dos furos-de-gusa são realizadas a partir de experiência operacional com incertezas elevadas devido ao grau de subjetividade envolvido no processo decisório. O objetivo da estimação do nível dos líquidos do cadinho é a melhoria do processo de abertura e fechamento dos furos-de-gusa como forma de se estabilizar a estrutura física dos furos-de-gusa e temperaturas de parede, preservando os refratários da região interna do furo e evitar interferências na distribuição de gás e descida de carga no alto-forno (NIGHTINGALE; TANZIL, 2001).

3.3 Força-Eletromotriz Gerada na Carcaça do Alto-Forno

Um dos sistemas de avaliação do nível dos líquidos em altos-fornos que vem sendo utilizado em algumas plantas é a medição da força-eletromotriz gerada na carcaça do alto-forno, doravante chamada FEM. Porém, a medição e técnicas de interpretação deste sinal não evoluíram o suficiente para que o sistema fosse confiável o bastante para servir de base para uma otimização da operação e alterações na estratégia de esgotamento dos altos-fornos (PETERS, 2001).

As FEMs são produzidas devido aos processos químicos e térmicos que ocorrem no forno. Estes processos geram células elétricas na região devido a oxidação do ferro para formar o óxido de ferro e escória. A corrente produzida por estas células migra para a carcaça do forno através de zonas de diferentes resistências e gradientes de calor. A FEM

medida na carcaça do alto-forno é indicativo do estado térmico e químico no interior do forno e também do nível dos líquidos no cadinho (RADILOV, 1985), (DOROFEEV; NOVOKHATSKII, 1984), (DUBOVIK, 1982) e (RUFF, 1927).

Como pode ser visualizado na Figura 22 (DOROFEEV; NOVOKHATSKII, 1984), para um processo estável, a corrente e, conseqüentemente, a diferença de potencial que é medida na carcaça (FEM) varia de acordo com três componentes: a resistência da camada de coque (R_{Co}), a resistência da camada de ferro-gusa (R_G) e a escória (R_E). Depois de esgotado, o cadinho é preenchido em sua grande parte por coque. Quando o nível de líquido no cadinho aumenta, o ferro e a escória fluem para o aglomerado de coque preenchendo os espaços entre eles. Isto causa uma elevação na resistência interna, principalmente devido à escória, e um correspondente aumento da FEM medida. Quando o alto-forno é esgotado, os níveis dos líquidos caem, a resistência interna cai e a FEM medida decresce.

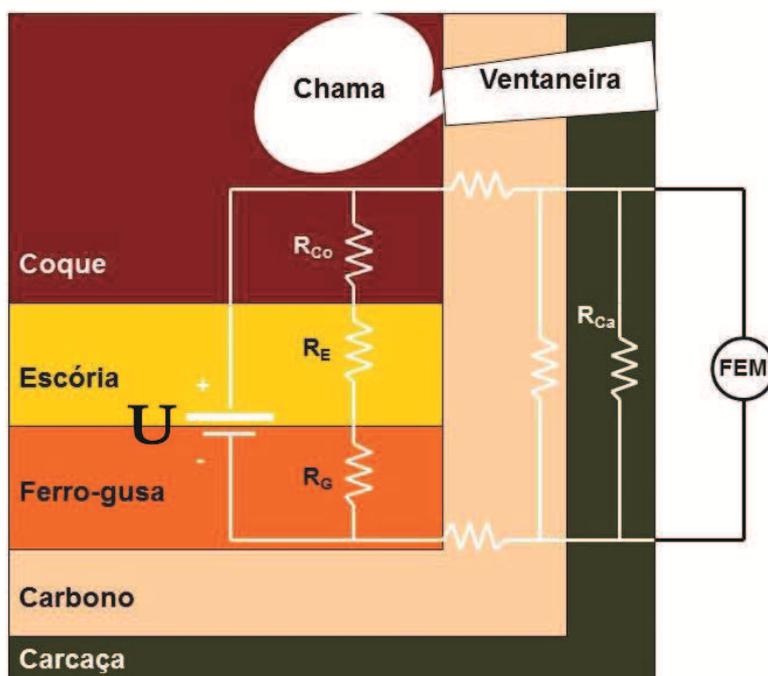


Figura 22: Circuito elétrico equivalente à geração de FEM na carcaça de um alto-forno.

Com o alto-forno em uma dada condição operacional constante (parâmetros de sopro, limites químicos das matérias-primas, etc...), U deve permanecer relativamente constante e pode ser considerada como a FEM medida no final do esgotamento do cadinho (cadinho vazio). Mudanças neste valor são indicativos de mudanças na temperatura do metal e composição (particularmente silício e manganês). A tensão elétrica U irá mudar com o estado térmico do alto-forno, mas isto não é notado como uma alteração obrigatória de R_{Co} . Contudo, a mudança do estado térmico do cadinho deve ser considerada quando for usado o sinal FEM para detecção de nível de líquidos no cadinho do alto-forno pois podem afetar a FEM medida (DOROFEEV; NOVOKHATSKII, 1984).

Alguns estudos como (DUBOVIK, 1982) e (BRAMMING; HALLIN, 1989) mostraram que a monitoração da variação da diferença entre FEMs por meio de 2 sensores, um sobre o nível do anel de vento e outro na região inferior do cadinho possui boa correlação com o nível dos líquidos no cadinho. A FEM medida mostra uma variação de longo prazo no nível mínimo (fim do esgotamento) e o nível máximo (início do esgotamento). Esta

variação pode estar relacionada com variações na drenagem do cadinho ou pode ser o resultado de mudanças nos níveis térmicos do cadinho.

O sinal produzido pela FEM gerada entre os pontos inferior e superior do cadinho é muito pequeno (variações menores que 1 milivolt) e consiste num somatório de 3 componentes: básica, periódica e ruído (alta frequência). A primeira tem relação com o estado térmico do cadinho, a diferença de temperatura entre os pontos de medição e também com os componentes contidos no gusa e escória. A componente periódica é associada ao enchimento e esgotamento dos líquidos no cadinho. A última, por sua vez, é consequência do movimento de materiais no forno e deve ser filtrada do sistema (PRONIN, 1985).

É de conhecimento que alterações nos parâmetros do processo tais como vazão, pressão ou temperatura de sopro também afetam a FEM pois impactam diretamente na catalisação do processo de redução (LEBED et al., 1990).

Análises da FEM medida e das temperaturas nos pontos de medição da carcaça indicam claramente que as variações na FEM e a sequência de esgotamento possuem frequências muito mais altas que as flutuações das temperaturas dos pontos de medição. A evolução da FEM não é correlacionada diretamente com a diferença de temperatura entre os pontos de medição da FEM. Contudo, esta temperatura pode ser utilizada para realização da correção do sinal puro da FEM medido pelos sensores, especialmente em caso de variações no estado térmico do cadinho (MIELENZ; KRUNER; KOCHNER, 2008).

Deve-se considerar, também, a interferência devido à diferença de temperatura entre os pontos de conexão na carcaça e os cabos de medição, conhecido como efeito termoelétrico ou efeito Seebeck, no caso em que os equipamentos sejam feitos de materiais diferentes. Cabos de cobre, em especial, geram tensões termoelétricas excessivas e não devem ser utilizados para medições de FEM na carcaça do cadinho. A tensão termoelétrica gerada em junções aço/cobre é da ordem de $7\mu V/^{\circ}C$. Variações de 12 graus Celsius geram tensões próximas de $100\mu V$ entre os sensores da FEM, ocasionando desvios significativos nos valores medidos. Sendo assim, recomenda-se a utilização de cabos baseados em ferro carbono pois outro material potencializaria este efeito, dificultando a análise (PETERS, 2001).

Os sensores instalados na carcaça e sensores introduzidos nos blocos de carbono (camada refratária existente na região interna do cadinho) apresentam o mesmo comportamento, porém, a FEM medida no sensor do bloco de carbono apresenta maior amplitude (PETERS, 2001). Neste artigo também se relata que, por tratar-se de medição elétrica de uma tensão muito pequena, podem ocorrer interferências elétricas tais como curtos-circuitos à terra de equipamentos elétricos próximos, máquinas de solda, entre outros. Contudo, essas perturbações não são muito comuns e são de fácil identificação.

Correlações entre a concentração de silício e a FEM apresentadas em (LEBED et al., 1990), sugerem que o estado térmico do cadinho é diretamente proporcional ao valor FEM, ou seja, se o silício cai, a FEM medida também cai pois mudanças nas concentrações de Si no ferro-gusa são indicativos de alterações no estado térmico do alto-forno.

Uma boa relação entre a temperatura do gusa e a FEM pode ser obtida, mas esta relação não é sempre verificada. Além disso, diferentes observações em processos industriais confirmaram que a série temporal da FEM possui variância não constante, ou seja, não é um sinal estacionário (LEBED et al., 1990) e (MIELENZ; KRUNER; KOCHNER, 2008). Seu comportamento deve-se a dois fenômenos distintos, que ocorrem em duas faixas de frequências também distintas:

1. Sinal de alta-frequência que tem relação com a variação da resistência do circuito elétrico induzido pelas variações dos líquidos e suas quantidades. Este comporta-

mento deve-se apenas às alterações das resistências R_G e R_E (ver Figura 22) que, por sua vez, dependem da altura da camada de líquidos (nível);

2. Sinal de baixa-freqüência que tem relação com as modificações no estado global do cadinho (permeabilidade, carga, temperatura, etc.). Este comportamento deve-se a alterações da resistência R_{Co} (ver Figura 22), e também a velocidade de produção de elétrons na fonte geradora (ritmo operacional).

A Figura 23 (MIELENZ; KRUNER; KOCHNER, 2008) mostra que a FEM tem boa correlação com a seqüência de esgotamento e com o nível do líquido no cadinho. Nesta figura, observa-se que durante o tempo entre duas corridas, o sinal FEM aumenta, indicando que o cadinho está enchendo. A FEM aumenta com o começo da corrida e decresce somente quando a escória começa a fluir do cadinho. Isto está relacionado com baixas taxas de esgotamento ($<4t/min$) sugerindo que a taxa de chegada do líquido no cadinho é maior que a taxa de extração. O início da saída da escória corresponde ao maior nível da FEM, por que reflete a maior resistência interna no circuito elétrico devido à camada de escória está no máximo e sua resistência ser muito elevada neste momento (LEBED et al., 1990).

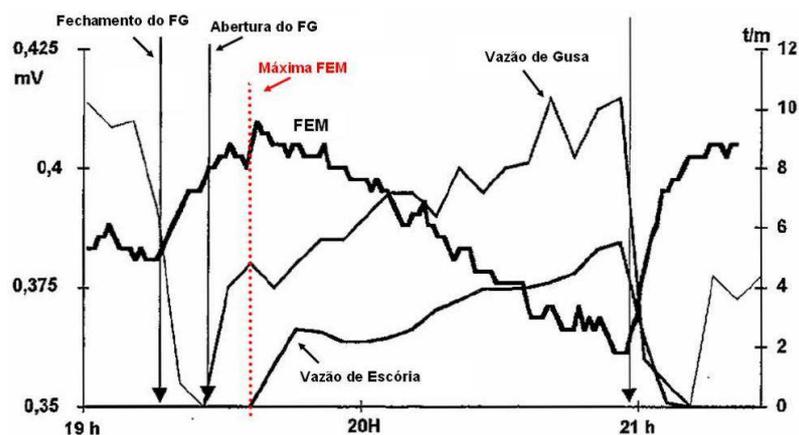


Figura 23: Comportamento da FEM medida em relação ao esgotamento dos líquidos no cadinho.

Ainda na Figura 23, no início da corrida, o diâmetro do furo-de-gusa corresponde ao diâmetro da broca utilizada na perfuratriz. Por isso, a vazão de esgotamento é pequena levando a uma continuidade do processo de enchimento do cadinho (taxa de produção de líquido maior que a taxa de extração). Durante a corrida, a área do furo-de-gusa se torna maior, ocasionando uma elevação na vazão de esgotamento. Com este acréscimo na vazão, o nível de enchimento do cadinho alcança o máximo e, tão logo a taxa de extração supere a taxa de produção, o nível cai.

Não foi encontrada, até o momento, nenhuma dependência clara entre alterações de parâmetros operacionais do alto-forno e a FEM. Contudo, o sinal parece ser influenciado pelas condições termoquímicas do cadinho. Um método eficiente de se contornar o problema é criar uma estratégia de abertura dos furos-de-gusa com base apenas na análise da variação da FEM no intervalo entre poucas corridas (considerando apenas a componente de alta-freqüência do sinal). Outro método efetivo é a compensação da variação de longo prazo com a utilização da técnica de médias móveis temporais ou filtros passa-alta, resultando um sinal muito estável a longo prazo (MIELENZ; KRUNER; KOCHNER, 2008).

As variações da FEM em função da sequência de abertura dos furos-de-gusa têm frequência muito maior que as flutuações de temperatura nos pontos de contato. A evolução da FEM durante o esgotamento não está relacionado diretamente com a diferença de temperatura dos pontos de contato. Contudo, a informação desta temperatura pode ser utilizada para realizar a correção da FEM medida nos sensores, especialmente durante mudanças no estado térmico do cadinho (LEBED et al., 1990).

3.4 A Relação entre o Processo de Esgotamento do Cadinho e a FEM

A medição da FEM é uma ferramenta importante não somente para a medição no nível de enchimento do cadinho, mas também para a detecção de condições críticas e ajustes na estratégia de abertura dos furos-de-gusa de acordo com as condições do cadinho.

Uma informação muito importante para uma boa estratégia de abertura dos furos-de-gusa do alto-forno é o ponto de mudança de inclinação na curva da FEM: se o ponto de mudança não ocorre, o nível do líquido continuará subindo e uma abertura paralela dos furos-de-gusa deve ser realizada. Sendo assim, a FEM pode ser utilizada para ajudar na detecção de problemas de esgotamento do cadinho como podemos visualizar na Figura 24 (PETERS, 2001).

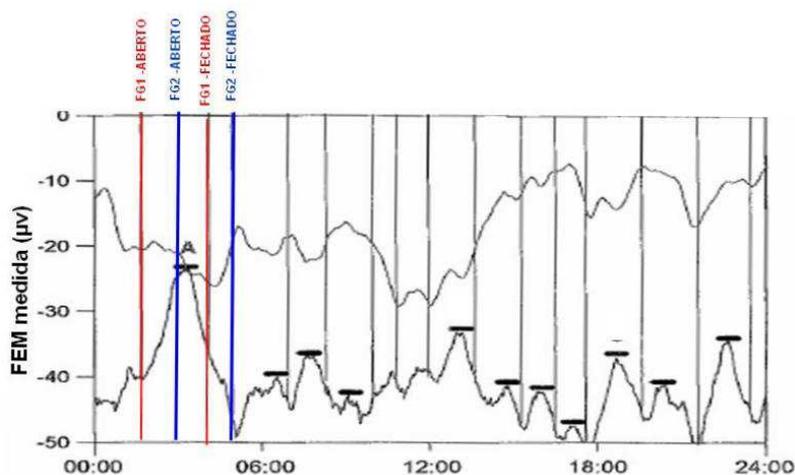


Figura 24: Identificação de problemas de esgotamento do cadinho pela FEM.

Nota-se claramente acima que, no pico A, o sinal FEM continuou subindo, mesmo após abertura do furo-de-gusa #1, indicando uma retenção anormal de líquidos. As ações corretivas foram a abertura do segundo furo-de-gusa em paralelo (furo-de-gusa #2) e a redução brusca da pressão de sopro durante a corrida permitindo uma maior vazão de extração (redução da interface de escória e o volume de escória no furo-de-gusa). Após estas ações, nota-se que nas corridas posteriores, o nível dos líquidos permaneceu oscilando em patamares normais de operação (PETERS, 2001).

Pela utilização da FEM, foi verificado que corridas paralelas (dois furos-de-gusa abertos ao mesmo tempo) de longa duração ou corridas com tempos reduzidos induzem fortes variações nos níveis de líquidos do cadinho. No entanto, a meta operacional é conseguir escoamento constante de metal líquido e escória, sem fortes variações nos níveis de líquido. Esta meta foi alcançada em um alto-forno na Alemanha, reduzindo-se a velocidade de vazamento e encurtando-se o tempo de corridas paralelas. Isso, por sua vez, levou a uma redução no número de corridas por dia para 7 ou 8 e prolongou-se o tempo

de escoamento de escória para quase 98% do tempo total de corrida. A suavização do escoamento de gusa e escória, bem como a redução da velocidade de escoamento, levou a uma redução no desgaste do forro do cadinho. Além disso, a quantidade de perfurações e quantidades de material para fechamento dos furos-de-gusa foram reduzidos significativamente (MIELENZ; KRUNER; KOCHNER, 2008).

Com a utilização de medições de vazão de saída de ferro-gusa e escória do alto-forno (medição de nível no carro torpedo e medição de pressão no sistema hidráulico do tambor de granulação, respectivamente) e taxas de produção instantâneas (baseados na análise de gás do topo e parâmetros de sopro), foi desenvolvido em (BRANNBACKA; SAXEN, 2004) um modelo caixa-branca (baseado em equações físico-químicas do processo) para estimação dos níveis de gusa e escória no cadinho do alto-forno. Este trabalho ressalta que a quantidade de ferro-gusa encontrada no cadinho tem forte correlação com a FEM medida na carcaça do alto-forno.

3.5 Instalação do Sistema de Medição do Sinal FEM

Os equipamentos que compõem o sistema de medição da força-eletromotriz na carcaça do alto-forno foram instalados no alto-forno 1 da ArcelorMittal Tubarão em Vitória no Espírito Santo para avaliação do comportamento deste sinal e se o mesmo estava de acordo com a literatura especializada avaliada. Com base nas informações obtidas através dos artigos técnicos sobre o assunto, foram instalados 4 conjuntos de sensores na carcaça do alto-forno, um conjunto (par) por furo-de-gusa. Estes sensores medem a diferença entre as forças-eletromotrices geradas na carcaça das regiões inferiores (fundo do cadinho) e superiores (acima dos níveis das ventaneiras), conforme pode ser visualizado na Figura 25 e em acordo com a metodologia desenvolvida em (GOMES, 2010).

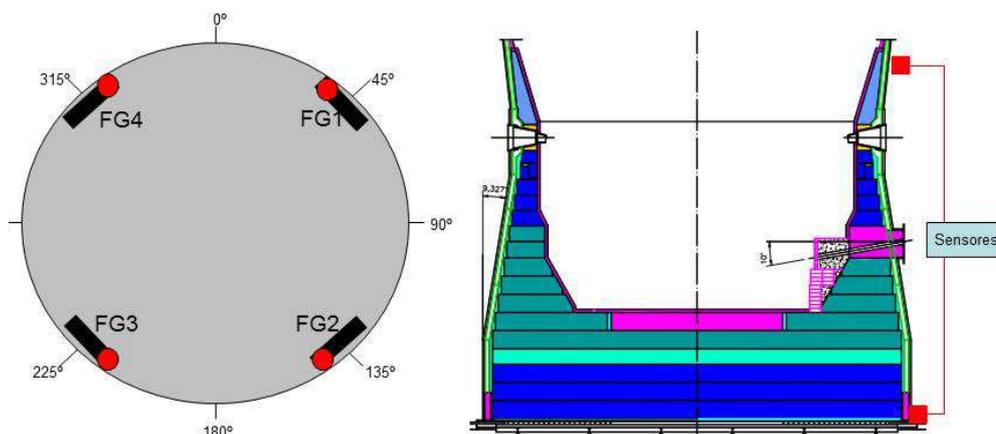


Figura 25: Disposição Física dos Sensores FEM do Alto-Forno 1 (Vista Superior e de Perfil).

Com todo o sistema integrado e, após alguns ajustes, foram encontrados sinais que apresentavam correspondência às descrições encontradas na literatura especializada, ou seja, sinais que aumentam quando os furos-de-gusa estão fechados e diminuem quando estes se encontram abertos, conforme pode ser visto na Figura 26. Nesta figura, quando o furo-de-gusa 4 é fechado (linha preta igual a zero) no tempo 160 minutos, por exemplo, o sinal da FEM (linha azul) aumenta rapidamente e só inverte o sentido 20 minutos após o furo-de-gusa 1 ser aberto no tempo 170 minutos. Este atraso acontece pois os furos-

de-gusa são abertos com brocas pequenas (66mm), iniciando a extração do material com baixo fluxo para garantir a segurança da operação e bom esgotamento do cadinho. Depois de algum tempo (20 minutos neste caso) o furo começa a aumentar, devido ao ataque da escória e gusa, fazendo o fluxo de saída de material pelo furo ser maior que o fluxo de material que entra (produção do alto-forno - AF) no cadinho. Neste instante, o sinal FEM inverte o sentido, indicando que o nível de líquidos no cadinho está diminuindo. Nesta figura, a linha azul clara representa a pressão de óleo no tambor do granulador de escória, indicando a quantidade de escória que está saindo do alto-forno.

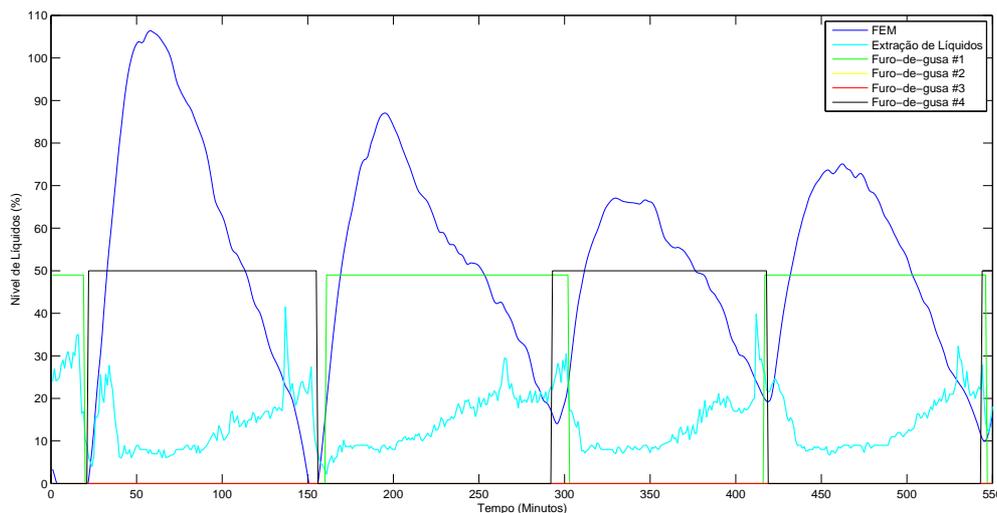


Figura 26: Comportamento da FEM medida na carcaça do Alto-Forno 1.

Ainda, na Figura 26, notamos que o sinal (média dos 4 conjuntos de campo) apresenta-se como sendo, aparentemente, não-estacionário. Isto deve-se à sua componente correlacionada ao estado térmico e químico do cadinho (PRONIN, 1985) e (MIELENZ; KRUNER; KOCHNER, 2008), à boa correspondência entre aberturas e fechamentos dos furos-de-gusa e a vazão de extração de escória com o comportamento de subida e descida da FEM, mais uma vez em consonância com a literatura especializada.

Este sinal foi disponibilizado de imediato para uso dos operadores da planta pois já apresenta informações importantes do processo de esgotamento do cadinho. Outro dado interessante é que, como os 4 (quatro) sinais obtidos apresentam similaridade, foi adotada a média destes sinais para utilização neste trabalho. Contudo, os sinais de cada conjunto de medição FEM foi mantido na tela de operação conforme pode ser visualizado na Figura 27 pois estas informações são importantes para visualização de desbalanceamento no esgotamento entre canais, já que estes operam geralmente em pares.

Com base na relação causa/efeito que a FEM apresenta em relação aos furos-de-gusa, a modelagem por um sistema entrada/saída (modelo ARMAX) parece, num primeiro momento, bastante razoável. Contudo, os resultados nesta direção não foram satisfatórios devido ao grande número de variáveis que impactam direta e indiretamente neste sinal medido na carcaça do alto-forno, como por exemplo, temperatura e pressão no cadinho. Este elevado número de variáveis produz uma dissociação com variáveis que deveriam ter elevadas correlações (GOMES; SALLES, 2011).

Durante este projeto, as simulações e análises foram realizadas com o auxílio do software de simulação computacional MatLab 8.1 da empresa MathWorks Inc.

Para utilização do sinal EMF no desenvolvimento de todo o trabalho a posterior,

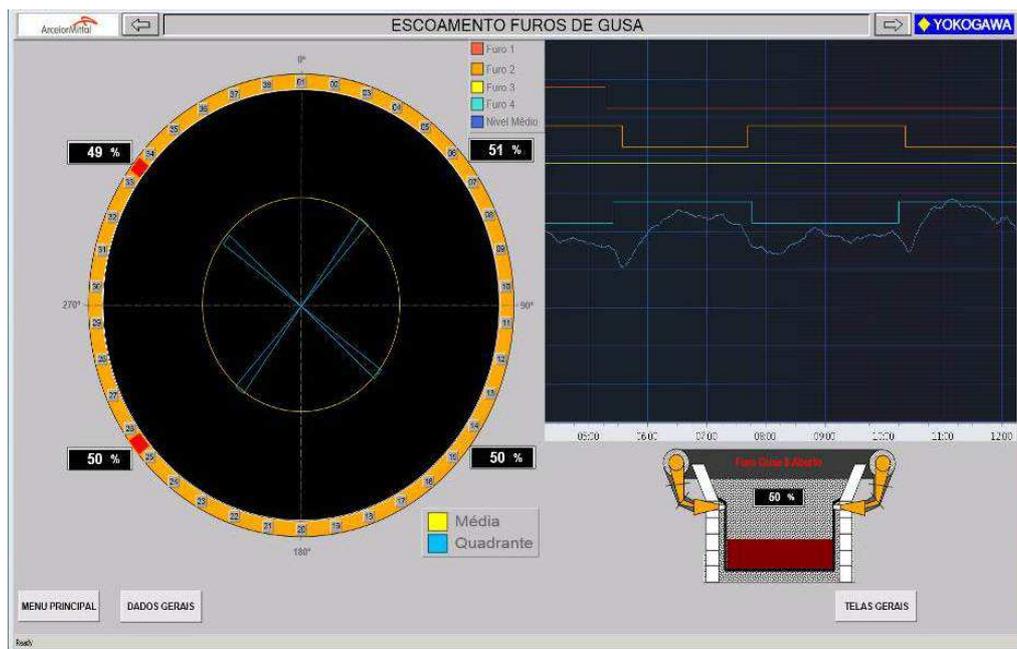


Figura 27: Tela do Nível dos Líquidos FEM do Alto-forno 1.

optou-se pela utilização da taxa de amostragem $t = 5$ minutos para um período médio entre aberturas de furos-de-gusa de 155 minutos. Desse modo, trabalhou-se com a relação de frequências de 31 vezes (período do sistema / amostragem) ficando acima do intervalo usualmente indicado mas que justifica-se pela atualização mais rápida da previsão que será realizada mais adiante no trabalho. Lembrando que, pelo teorema de Nyquist, para que um sinal amostrado retenha as características fundamentais do sinal original, é necessário que o tempo de amostragem seja suficientemente curto. Desta forma, um sinal deve ser amostrado com taxa de amostragem de, pelo menos, duas vezes maior que a constante de tempo do processo em análise. Na prática, entretanto, a frequência de amostragem é normalmente escolhida entre 5 e 10 vezes maior do que a maior frequência de interesse contida nos dados, e não apenas 2 vezes maior (chamada frequência de Nyquist), como exigido por este teorema (AGUIRRE, 2007).

3.6 Filtragem do Sinal FEM

Como citado na subseção 3.3, o comportamento do sinal da FEM gerada na carcaça do alto-forno é afetado por diversos fenômenos tais como: nível dos líquidos, estado termoquímico do cadinho, estado do homem-morto, etc. Estes efeitos tornam-se fortemente indesejados para o desenvolvimento do trabalho já que dificultam a análise comparativa entre esgotamentos atuais e passados, impactando na modelagem por meio de séries temporais sem utilização de transformações matemáticas mais complexas. Diante disto, para realizarmos a separação do sinal de nível dos líquidos no interior do cadinho do sinal da FEM torna-se necessária a caracterização detalhada de todos os componentes deste sinal com a ajuda das análises temporal e de frequências.

Pela análise da FEM no tempo em um espaço temporal de aproximadamente 7 dias, como pode ser visualizado na Figura 28, foram identificados dois comportamentos distintos: o primeiro, de menor período, está fortemente relacionado com o processo de abertura e fechamento dos furos-de-gusa, e o segundo, de maior período, está relacionado

com o estado térmico-químico do cadinho (DUBOVIK, 1982) e (BRAMMING; HAL-LIN, 1989).

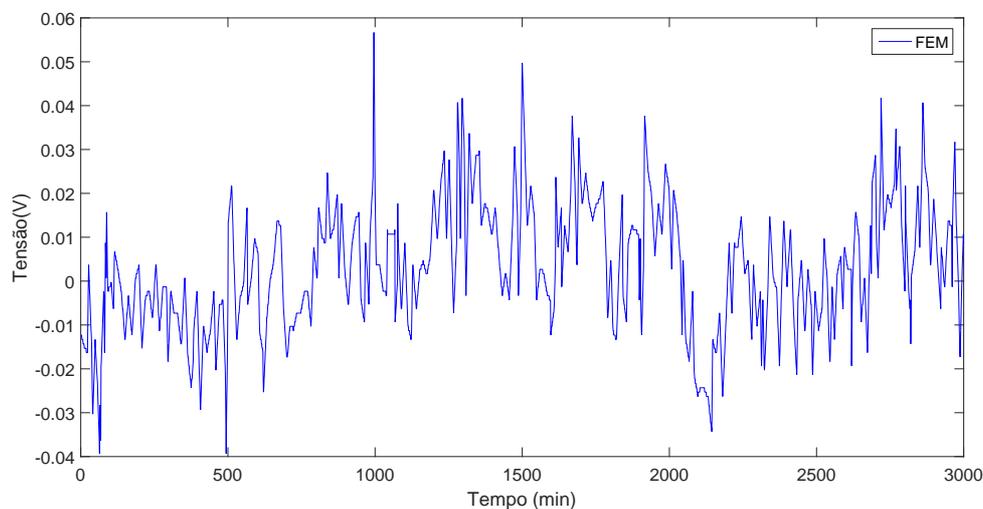


Figura 28: Análise Temporal da FEM.

Pela análise do módulo da transformada rápida de Fourier (KUO, 1995), identificou-se que existem algumas regiões de frequências dominantes no sinal, mostradas na Figura 29. De uma forma geral, as frequências muito baixas (muito próximas de zero) estão relacionadas com a média do sinal e devem ser descartadas, as frequências encontradas na região central estão relacionadas com as aberturas e fechamentos dos furos-de-gusa e por fim, as frequências mais altas possuem densidades espectrais bem mais distribuídas e estão relacionadas, principalmente, aos ruídos de medição. Desta figura, podemos notar uma maior concentração de energia entre $66,7\mu$ e $166,7\mu$ Hertz, que correspondem aos períodos de 250 a 100 minutos, respectivamente (alto-forno estável e com ritmo de produção nominal).

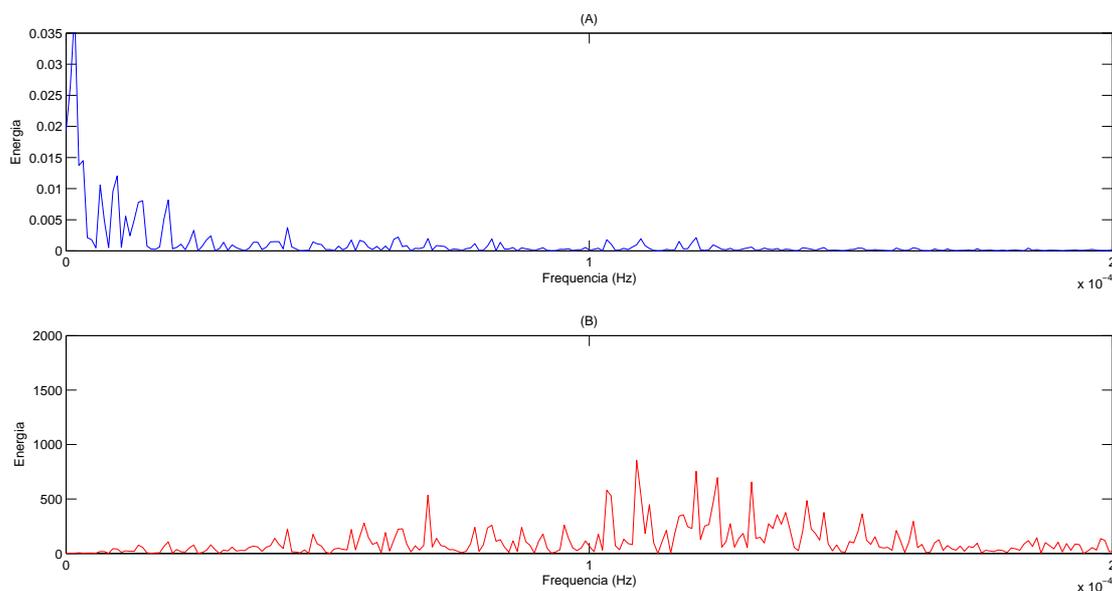


Figura 29: Transformada de Fourier da FEM e do Nível de Líquidos.

Analisando-se o módulo da transformada de Fourier com foco maior na região de interesse, verificou-se mais uma vez duas bandas de frequências predominantes: baixas frequências (períodos entre 20 e 60 horas) e altas frequências (períodos entre 100 e 180 minutos). As primeiras, possivelmente, representam o estado termoquímico do cadinho e valores médios (nível dc do sinal) e as últimas, estão relacionadas com o processo de abertura e fechamento dos furos-de-gusa.

Com base nas análises temporal e de frequências, realizou-se a filtragem digital do sinal centralizada nas frequências de interesse (aberturas e fechamentos dos furos-de-gusa, entre 100 e 250 minutos). Esta filtragem foi realizada no sistema de controle através da associação de filtros passa-baixas e passa-altas de 1^a ordem do tipo Butterworth resultando em um sinal com energia concentrada nesta região e com atenuações significativas nas frequências fora dos limites dos filtros. Após esta filtragem, obtivemos um sinal, aparentemente, estacionário e altamente independente do estado termoquímico do cadinho (oscilações de baixa frequência) e que concentra energia na região de frequências de abertura e fechamento dos furos-de-gusa, conforme pode ser visto na Figura 30.

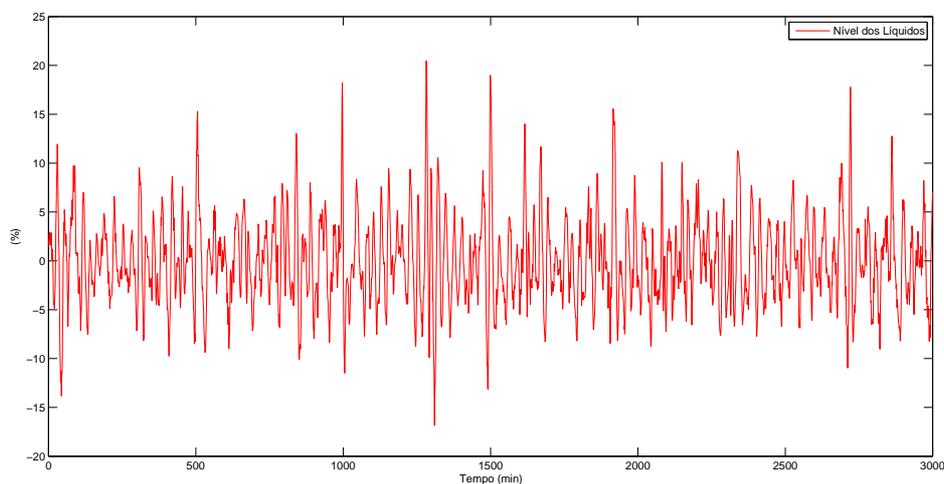


Figura 30: Análise Temporal do Nível de Líquidos.

Nesta figura o sinal da FEM é dado em percentual, onde o zero indica o nível médio deste sinal. Quando o sinal da FEM for dado em percentual será chamado de nível ao longo do texto. O sinal filtrado, também revela alguns fenômenos interessantes como um maior detalhamento do comportamento dos líquidos no interior do cadinho e mudanças no interior do cadinho causadas por alterações no ritmo operacional ou de matérias-primas do alto-forno (GOMES; SALLES, 2011).

4 MODELAGEM E PREVISÃO DO NÍVEL DE LÍQUIDOS

O desenvolvimento de uma ferramenta que auxilie na tomada de decisão sobre o instante de abertura e fechamento dos furos-de-gusa do alto-forno pode contribuir para tornar esta operação mais confiável e assertiva, trazendo, conseqüentemente, a possibilidade de otimização deste processo com ganhos financeiros associados.

Os procedimentos apresentados a seguir constituem a junção de vários modelos e metodologias descritos nas literaturas de referência, incluindo adaptações para melhor representar o processo do alto-forno. Esses modelos retratam as principais dinâmicas do nível dos líquidos no interior do cadinho do alto-forno, podendo ser utilizados na rotina diária de operação (medição e predição) além da utilização como entrada em sistemas especialistas visando à otimização ou sintonia dos processos operacionais.

Como estudo de caso, utilizamos o Alto-Forno 1 da ArcelorMittal Tubarão (antiga Companhia Siderúrgica de Tubarão - CST), localizada na região da Grande Vitória, Espírito Santo. Este alto-forno é considerado um equipamento de alta produtividade possuindo 4 furos-de-gusa, volume interno de $4540m^3$, 38 ventaneiras, cadinho com 14,1 metros de diâmetro e $717m^3$ de volume, apresentando uma produção diária média de $9.600t/dia$ de gusa e $2700t/dia$ de escória. Na Figura 31 é apresentada uma visão panorâmica desta planta.



Figura 31: Visão panorâmica do Alto-Forno #1 da ArcelorMittal Tubarão.

O trabalho de elaboração da ferramenta preditiva para monitoração do nível dos líquidos no interior de cadinho de altos fornos é constituído pelas seguintes etapas: instalação

do sistema de medição da FEM (atualização do trabalho realizado em (GOMES, 2010)), identificação, caracterização e tratamento do sinal de nível dos líquidos obtido através da FEM, avaliação e comparação entre modelos, validação do modelo selecionado e implantação de sistema de previsão do nível dos líquidos no processo real produtivo. Dois tipos de filtros do sistemas de previsão serão apresentados neste capítulo. O primeiro utiliza o filtro de primeira ordem tipo Butheworth, implantado no Sistema Digital de Controle Distribuído (SDCD). O outro é um filtro FIR passa-faixa tipo janela de Kaiser desenvolvido usando o Matlab

4.1 Filtragem através do Sistema Digital de Controle

De acordo com o que foi definido no Capítulo 3, é necessária uma análise detalhada das características da série temporal visando escolher um modelo mais apropriado para modelagem e posteriormente, previsão. Esta avaliação consiste na análise de estacionaridade, linearidade e quando possível, um detalhamento do espectro de frequência da série temporal para detecção de sazonalidades.

A primeira análise consiste em verificar se o sinal filtrado pelo sistema digital de controle com o filtro de 1ª ordem é uma série estacionária. O teste de KPSS (Kwiatkowski, Phillips, Schmidt, e Shin) (KWIATKOWSKI et al., 1982) foi realizado no sinal filtrado e na FEM, apresentando os resultados contidos na Tabela 1. Estes resultados ratificam que a filtragem transformou o sinal em uma série temporal estacionária.

Tabela 1: Teste de Estacionaridade - Sinal Filtrado versus FEM

Sinal	KPSS	Hipótese nula
FEM	1	$h = 0$, série é trend-estacionária
Sinal Filtrado	0	$h = 1$, série é raiz unitária

A análise quanto à existência de não-linearidades nesta série foi realizada por meio dos testes de Lilliefors e Kruskal-Wallis. Para execução deste último, decidimos pela divisão do pacote de dados em cinco partes (10 dias de dados em cada parte) e foi testado para saber se o sinal tinha média e variância constantes. Os resultados podem ser visualizados na Tabela 2.

Tabela 2: Teste de Linearidade - Nível de Líquidos

Teste	Resultado	Possíveis Resultados
Lilliefors	1	$h = 0$, Os dados são normalmente distribuídos $h = 1$, Os dados não são normalmente distribuídos
Kruskal-Wallis	0,088	$p \ll 0$, Amostras não vem da mesma distribuição $p \gg 1$, Amostras vem da mesma distribuição

Na Figura 32 observa-se a distribuição visual dos pacotes em relação à média, variância, etc gerados no teste de Kruskal-Wallis. Com base nestas análises conclui-se que a série temporal possui componentes não-lineares.

Isto é confirmado pela Figura 33, onde se nota que o sinal apresenta um comportamento não linear concentrado nos valores próximos às bordas (níveis vazio e cheio dos líquidos). Nesta mesma figura, como o sinal analisado está filtrado, o eixo das abcissas

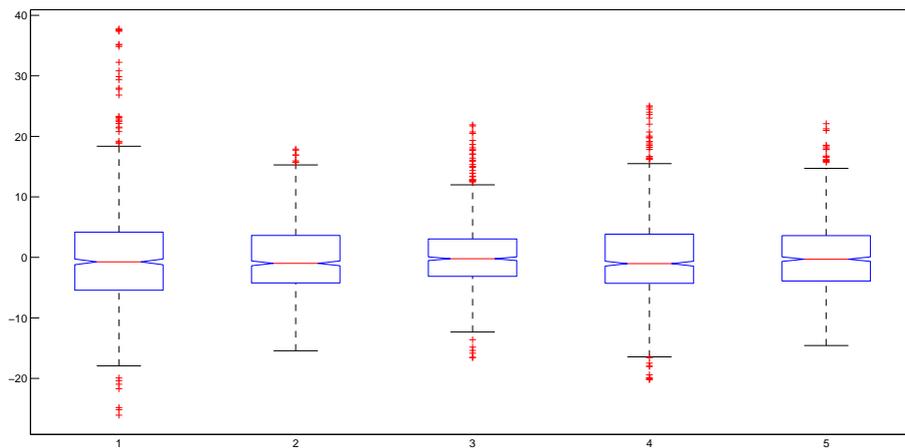


Figura 32: Comparação entre médias e variâncias de cinco pacotes de dados do Nível.

assume valores negativos devido a remoção do nível médio presente no sinal de baixa frequência.

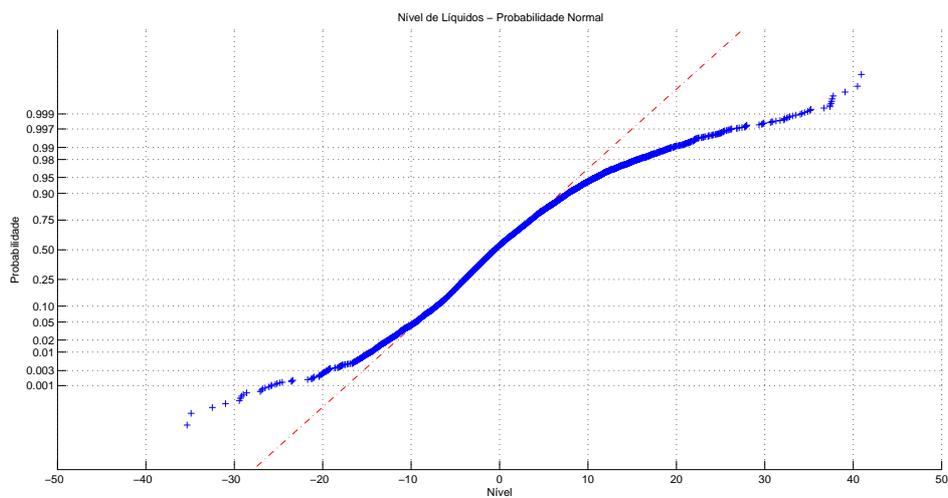


Figura 33: Sinal Filtrado de Nível dos Líquidos quanto à Linearidade.

As funções de autocorrelação e autocorrelação parcial do nível de líquidos apresentam decaimento muito lento e pouca variação em torno dos intervalos de confiança, respectivamente, conforme mostrado na Figura 34.

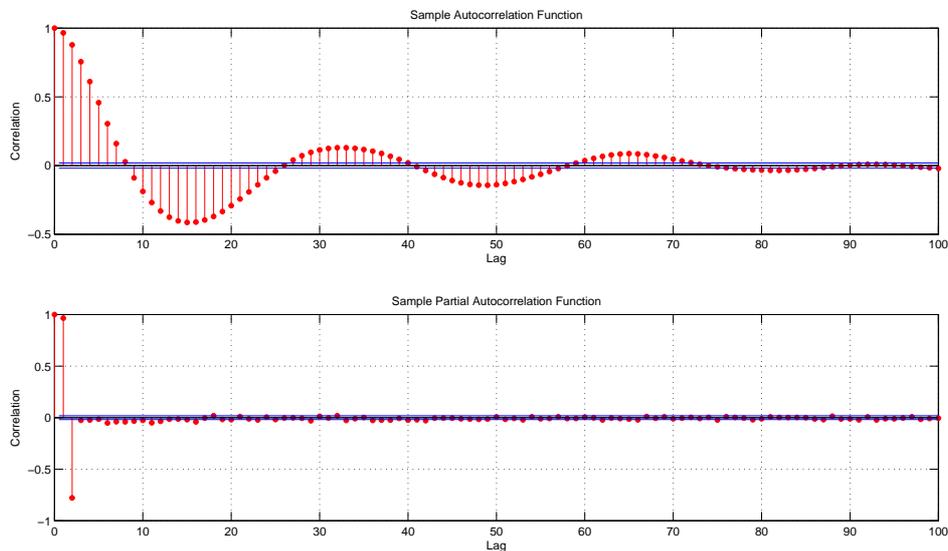


Figura 34: Autocorrelação e Autocorrelação Parcial do Nível de Líquidos.

Estes comportamentos associados são fortes sinais da existência de não-linearidades (G.KIRCHGASSNER; WOLTERS, 2007). Também pode ser notado, pelo comportamento da curva de autocorrelação e pela distribuição de energia do periodograma, que existem múltiplas sazonalidades (KARRA; KARIM, 2009) com causas inerentes ao processo de extração do ferro-gusa e escória do alto-forno. Foi escolhida a sazonalidade dominante (31 amostras (155 minutos)) para diferenciação do sinal para redução das oscilações na variância. Desse modo, o sinal foi diferenciado a cada 31 amostras, resultando nas ACF e PACF mostradas na Figura 35.

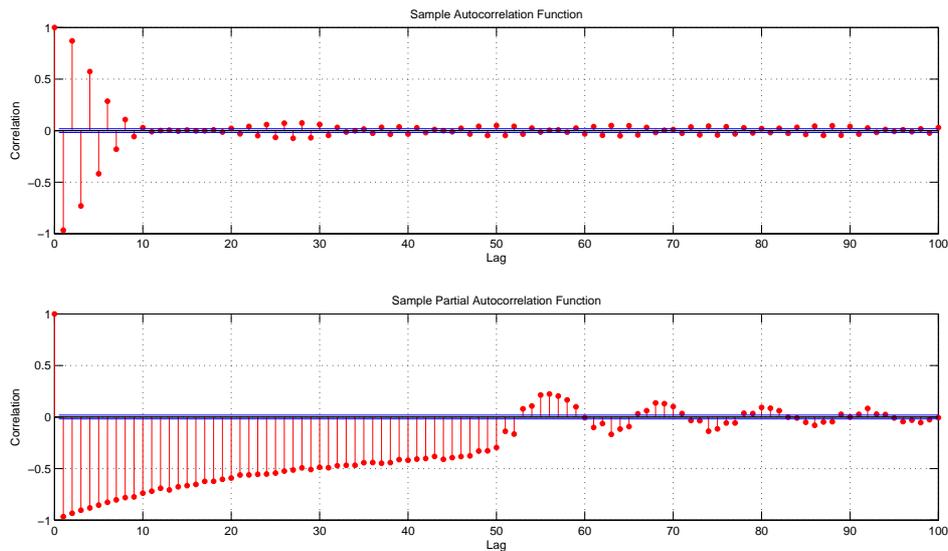


Figura 35: Autocorrelação e Autocorrelação Parcial do Nível de Líquidos Diferenciado.

Foi realizado novamente o teste de linearidade com este sinal diferenciado. Estes resultados estão na Tabela 3. Pode-se notar o resultado do teste de Lilliefors continuou negativo. Contudo, desta vez, o teste de Kruskal e Wallis apresentou resultado satisfatório ($p = 1$) conforme reforçado pela Figura 36. Este resultado implica que a média e variância do sinal podem ser considerados constantes, mas a não-linearidade ainda está presente.

Tabela 3: Teste de Linearidade - Nível de Líquidos Diferenciado

Teste	Resultado	Possíveis Resultados
Lilliefors	1	$h = 0$, Os dados são normalmente distribuídos $h = 1$, Os dados não são normalmente distribuídos
Kruskal-Wallis	1	$p \ll 0$, Amostras não vem da mesma distribuição $p \gg 1$, Amostras vem da mesma distribuição

Com o intuito de minimizar o efeito da não-linearidade neste sinal, foi aplicada a transformação de BoxCox (BOX; JENKINS; REINSEL, 1994) mas o sinal resultante manteve a não-linearidade das bordas como pode ser visto na Figura 37.

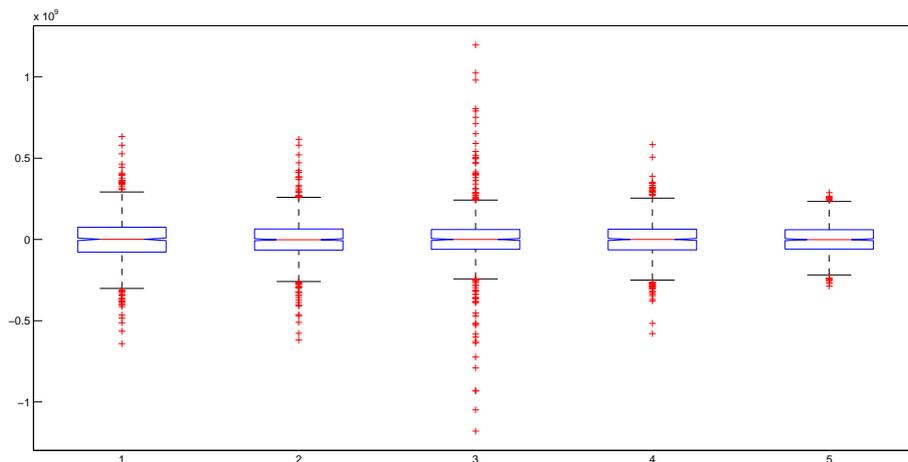


Figura 36: Comparação entre médias e variâncias de cinco pacotes de dados do Nível diferenciado.

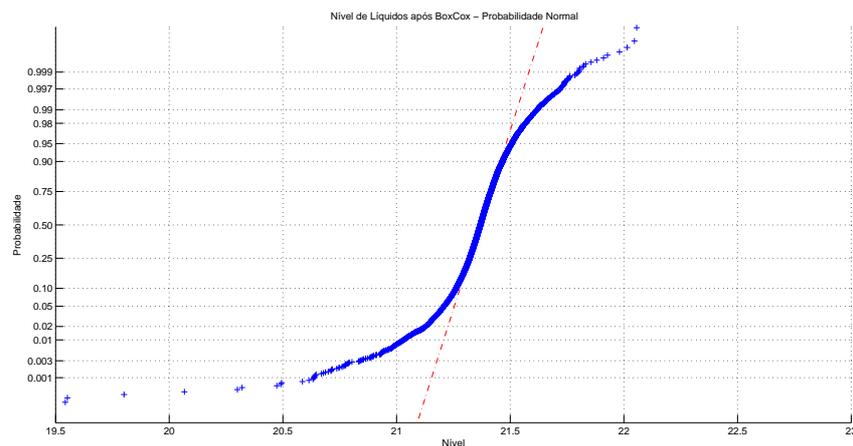


Figura 37: Sinal de Nível Diferenciado após Transformação de BoxCox.

4.2 Modelagem e Previsão do Nível de Líquidos

Neste ponto do projeto, sabe-se que o sinal filtrado e diferenciado é uma série temporal, estacionária e não-linear. A partir desta constatação, pode-se seguir dois caminhos para avaliar qual deles apresenta melhor performance para a previsão da série:

- Utilizar modelos lineares mesmo que não consigam representar todas as nuances do sistema ou;
- Desenvolver um modelo não-linear de menor criticidade de implantação no processo real.

4.2.1 Previsão com Modelos SARIMA Adaptativos

Técnicas de previsão são essenciais para garantir a segurança operacional de sistemas complexos. Contudo, em geral, não é fácil encontrar modelos precisos para descrevê-los. Recentemente, metodologias que utilizam séries temporais como ferramenta principal

têm atraído atenção crescente como uma aproximação eficiente para preencher estas lacunas. Em geral, sistemas reais podem ser modelados através de sistemas de múltiplas entradas e múltiplas saídas (*multi-input/multi-output* - MIMO) quando as conexões entre os sinais de entrada e saída são não-lineares, variantes no tempo e contaminados por ruídos (MATTHIAS et al., 1998).

Frequentemente, uma aproximação linear é aceitável quando o processo trabalha sem grandes variações dos pontos de operação. Porém, quando isto não é possível, métodos avançados de análise por séries temporais, como séries temporais adaptativas, podem ser usados como ferramentas eficazes para identificação e previsão em sistemas e fenômenos industriais reais (YANGMING et al., 2012), (ABO-HAMMOUR et al., 2012) e (GOLSHAN; SAMET, 2010).

Para modelagem deste sinal, foi utilizado o modelo SARIMA descrito na Equação 2.12. É importante observar que da análise de estacionariedade do sinal FEM filtrado, $d = 0$, $D = 1$ e $S = 31$. Desta forma, os modelos testados serão especificados a partir da classe $SARIMA(p, 0, q)(P, 1, Q)_{31}$. Como o nível dos líquidos diferenciado é um sinal não-linear, um modelo SARIMA com parâmetros invariantes no tempo pode realizar previsões satisfatórias apenas para curtos horizontes de tempo pois o erro de previsão aumenta significativamente para horizontes de previsão maiores. Este problema pode ser atenuado pelo uso de modelos SARIMA com parâmetros variantes no tempo (adaptativos).

A seleção da ordem do polinômios do modelo (p,q,P,Q) é obtida através da comparação dos erros de previsão dos modelos de diferentes ordens, checando-se a melhor performance através dos critérios BIC (*Bayesian Information Criterion*), AIC (*Akaike Information Criterion*) e MAE (*Mean Absolute Error*), como descrito em (BOX; JENKINS; REINSEL, 1994). Os resultados dos testes para identificação deste sinal são mostrados na Tabela 4.

Tabela 4: Modelos SARIMA(p,0,q)(P,1,Q)

Modelo	AIC (%)	BIC (%)	MAE(%)
(0, 0, 0)(1, 1, 0) ₃₁	201, 87	201, 58	824, 31
(0, 0, 0)(0, 1, 1) ₃₁	194, 09	193, 81	726, 88
(0, 0, 1)(0, 1, 1) ₃₁	161, 95	161, 75	388, 62
(0, 0, 2)(0, 1, 1) ₃₁	139, 14	138, 98	244, 86
(1, 0, 0)(0, 1, 1) ₃₁	124, 59	124, 44	175, 23
(2, 0, 0)(0, 1, 1) ₃₁	100, 49	100, 40	100, 73
(2, 0, 0)(0, 1, 2) ₃₁	100, 49	100, 41	100, 83
(3, 0, 0)(0, 1, 1) ₃₁	100, 48	100, 40	101, 10
(2, 0, 0)(1, 1, 1) ₃₁	100, 49	100, 41	100, 83
(2, 0, 1)(1, 1, 1) ₃₁	100, 47	100, 41	101, 10
(2, 0, 2)(1, 1, 1)₃₁	100, 00	100, 42	100, 00
(2, 0, 3)(1, 1, 1) ₃₁	100, 44	100, 42	101, 19
(3, 0, 1)(0, 1, 3) ₃₁	100, 02	100, 00	100, 37
(2, 0, 2)(2, 1, 1) ₃₁	100, 46	100, 44	101, 10

O uso do modelo $SARIMA(2, 0, 2)(1, 1, 1)_{31}$ apresenta os melhores resultados em dois critérios (AIC e MAE). Portanto, a equação do modelo para um horizonte de previsão

é mostrada abaixo:

$$\begin{aligned}\widehat{Z}_{k+1} = & +\phi_1 Z_k + \phi_2 Z_{k-1} + (1 + \Phi_1) Z_{k-30} - (\phi_1 + \phi_1 \Phi_1) Z_{k-31} \\ & - (\phi_2 + \phi_2 \Phi_1) Z_{k-32} - \phi_1 Z_{k-61} + \phi_1 \Phi_1 Z_{k-62} + \phi_2 \Phi_1 Z_{k-63} \\ & - \theta_1 a_k - \theta_2 a_{k-1} - \Theta_1 a_{k-30} + \theta_1 \Theta_1 a_{k-31} + \theta_2 \Theta_1 a_{k-32}\end{aligned}\quad (4.1)$$

onde o vetor de parâmetros $[\phi_1, \phi_2, \Phi_1, \theta_1, \theta_2, \Theta_1]$ deve ser escolhido através de um algoritmo recursivo focado na redução do erro total de previsão.

Um algoritmo genético foi utilizado para identificar os parâmetros do modelo, pois sua utilização não demanda cálculo do gradiente e não é suscetível a problemas de mínimos locais que surgem com superfícies de erro multimodais em comparação com o método da máxima verossimilhança (ML) (ONG; HUANG; TZENG, 2004), (CHEN; LEE; PENG, 2002) e (ELSAYED; SARKER; ESSAM, 2013). No entanto, os algoritmos genéticos, em geral, não garantem erro de estimação otimizado apresentando, muitas vezes resultados ligeiramente inferiores (YAO; SETHARES, 1994).

Um algoritmo genético recursivo chamado evolução diferencial (*Differential Evolution-DE*) é descrito em (STORN; PRICE, 1995). No algoritmo 1, a população de indivíduos é criada e submetida a operadores genéticos: seleção, mutação e recombinação (*crossover*). O melhor indivíduo da geração atual, que é aquele que apresenta a menor MAE de previsão durante o processo de avaliação dos modelos, é escolhido para ser comparado com os indivíduos da próxima geração até que o melhor indivíduo de todas as gerações é entregue como a solução do problema.

A função-custo usada para definir o melhor indivíduo é dada por $\sum_{k=N}^{2N} \|Z_{k+1} - \widehat{Z}_{k+1}\|/N$, onde N é o tamanho do pacote de dados e $\widehat{Z}_{k+1} = E[Z_{k+1} | Z_{k-N}, \dots, Z_k]$. Os principais detalhes do algoritmo DE são descritos abaixo:

O tamanho da população e a quantidade de parâmetros gerados são definidos para permitir que um número significativo de vetores-solução sejam testados dentro do mesmo ciclo (6.000 indivíduos) com uma demanda computacional aceitável. Nota-se que, já que o algoritmo garante que a função-custo da nova geração é igual ou inferior à geração anterior, alta mutação e reprodução são introduzidas (iguais a 0,9), permitindo uma ampla variação de indivíduos dentro do universo de soluções para elevar as chances de que o algoritmo consiga fugir de mínimos locais.

A previsão \widehat{Z}_{k+1} é dada através dos valores de passados e atual de Z_{k-N}, \dots, Z_k e a_{k-N}, \dots, a_k . Os indivíduos do algoritmo genético são definidos pelo vetor de parâmetros dos polinômios $(\phi, \theta, \Phi, \Theta)$ com suas respectivas ordens (p, q, P, Q) .

O algoritmo DE padrão foi modificado para garantir que $\Delta_S^{31} Z_k$ seja inversível e estacionário, ou seja, que $\Phi(B^S)\phi(B) = 0$, $\Theta(B^S)\theta(B) = 0$ e que tenham raízes fora do círculo unitário. Isto permite que o algoritmo calcule os resíduos a_k com base na seguinte equação:

$$a_k = \frac{\Phi(B^{31})\phi(B)\Delta_{31}^1 Z_k}{\Theta(B^{31})\theta(B)}, \quad (4.2)$$

Os parâmetros do modelo são calculados para a previsão do sinal de um passo a frente (5 min) através do algoritmo DE como mostrado na Equação (4.3), e os resultados de previsão são mostrados na Figura 38:

$$[\phi_1, \phi_2, \Phi_1, \theta_1, \theta_2, \Theta_1] = [1, 74, -0, 79, -0, 11, -0, 11, 0, 10, 0, 73] \quad (4.3)$$

```

Maxpop = 60 - Tamanho da população
Maxgen = 100 - Número de gerações
Mutaçã=0.9 - Taxa de mutação
Crossover = 0.9 - Taxa de reprodução crossover
Função-custo=menor MAE
Gera população inicial aleatória
Calcula valor da função-custo
Seleciona o individuo com melhor função-custo como vetor-alvo
while Numgen < Maxgen do
  Newpopulation = 0
  Numpop = 0
  while Numpop < Maxpop do
    Seleciona o vetor-base (população inicial)
    seleciona dois vetores aleatórios (população inicial)
    Gera mutação (vetores aleatórios e vetor-base)
    Cruza o indivíduo mutante com o vetor-alvo
    Soma o novo indivíduo à nova população
    Numpop = Numpop + 1
  end
  Calcula os valores da função-custo (nova população)
  Compara o vetor-alvo com a nova população
  Aplica os testes de inversibilidade e estacionaridade
  Seleciona o melhor indivíduo como vetor-alvo
  Numgen = Numgen + 1
  População inicial=Newpopulation
end
Melhor solução = Vetor-alvo

```

Algorithm 1: Evolução Diferencial

Este modelo apresenta 99,05% de precisão (percentual de vezes que o erro de previsão é menor que 10%), e o erro de previsão (MAE) é igual a 1,21 quando validado em um pacote de 6132 dados (equivalente a 21,29 dias). A autocorrelação dos resíduos e a correlação cruzada entre os resíduos e valores passados do nível dos líquidos podem ser visualizados nas Figuras 39 e 40, respectivamente.

Estas análises confirmam que há uma pequena correlação entre os resíduos causada por outras sazonalidades e não linearidades que estão presentes no sinal, mas decidiu-se não incluir mais parâmetros na expressão (4.1) pois a precisão de previsão não aumenta significativamente com a elevação de ordem; Além disso, isto dificultaria a implementação no software no sistema de controle digital do alto-forno (princípio da parsimonia).

O tempo de previsão de longo prazo foi definido com base na demanda de processo. O processo de abertura e fechamento dos furos-de-gusa, para este alto-forno, possui período médio de 155 minutos (31 *lags*) como apresentado em análise prévia. Além disso, os pontos mais importantes para monitoração são os níveis correspondentes a cadinho cheio e cadinho vazio. Portanto, decidiu-se pelo horizontes de 60 minutos à frente (quase a metade da sazonalidade, $h = 12$).

4.2.2 Previsão com Redes Neurais

Para utilização nesta série temporal, foram desenvolvidas 2 redes neurais que podem ser visualizadas nas Figuras 41 e 42. A diferença básica entre elas é que na segunda

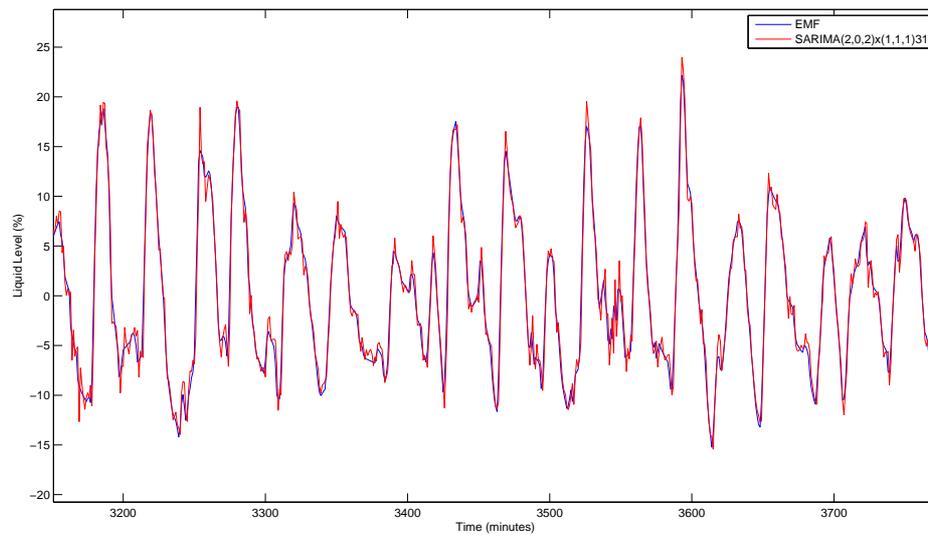


Figura 38: Previsão um Passo à Frente com Modelo SARIMA Adaptativo

foi inserida mais uma camada oculta e esta possui uma retroalimentação com a saída. O objetivo é tentar minimizar os erros de previsão com o aumento dos horizontes. Em ambas as redes foi realizado o treinamento supervisionado com base no pacote de 8000 dados (aproximadamente 28 dias).

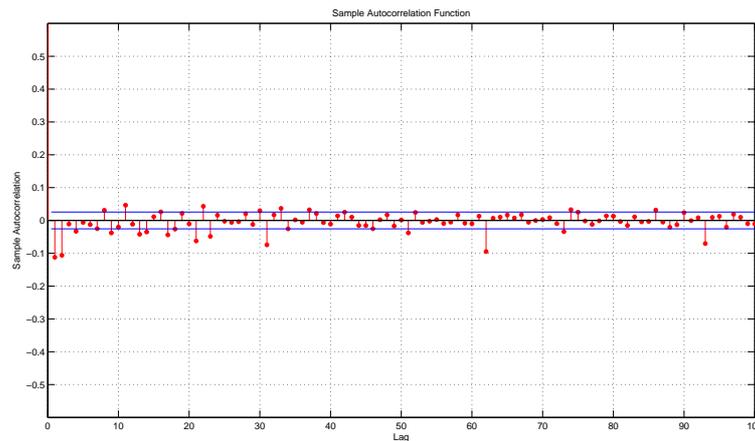


Figura 39: Auto-correlação dos Resíduos do Modelo SARIMA para Previsão de um Passo à Frente.

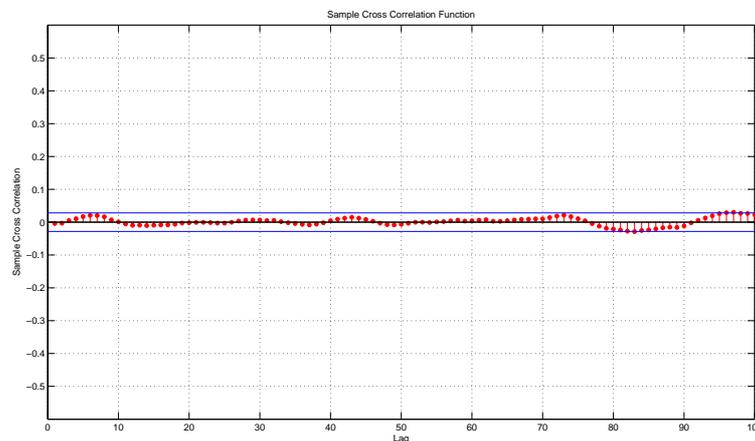


Figura 40: Correlação Cruzada entre os Resíduos do Modelo SARIMA para Previsão de um Passo à Frente e os Valores Passados de Z_k .

A primeira é uma RNA de atraso de tempo (TDNN) com os seguintes parâmetros:

- Duas camadas, sendo uma oculta;
- 100 neurônios na primeira e 1 neurônio da última camada;
- Funções de transferência tansig e purelin para a primeira e última camada, respectivamente;
- Atrasos de tempo de 0 a 11 amostras na entrada;
- Utilização de *bias* em todas as camadas.

A segunda é uma RNA recorrente (RNN) com os seguintes parâmetros:

- Três camadas, sendo duas ocultas;
- 100 neurônios na primeira e 1 neurônio nas outras camadas;

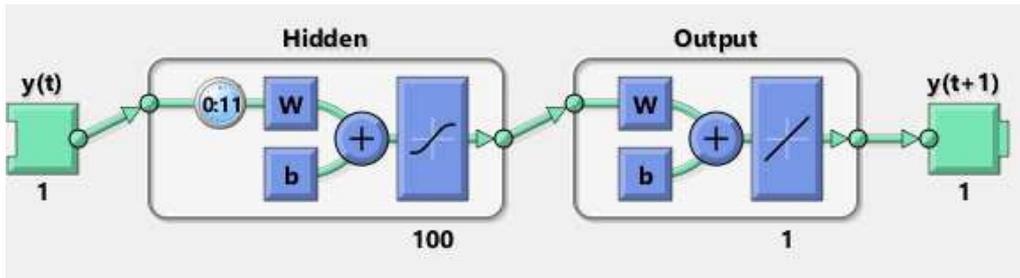


Figura 41: Rede Neural Artificial com Atraso.

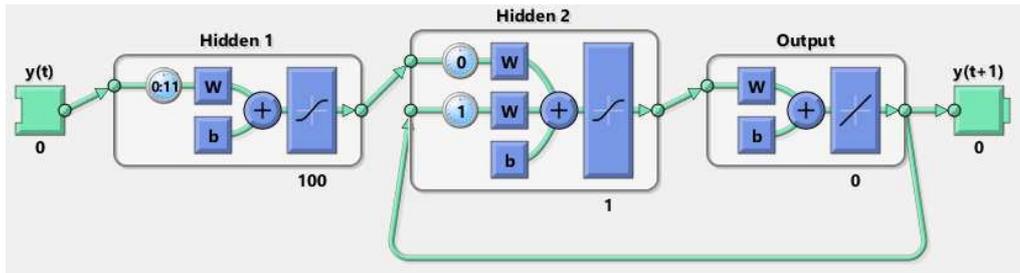


Figura 42: Rede Neural Artificial Recorrente.

- Funções de transferência tansig para a primeira e segunda e purelin na última camada;
- Atrasos de tempo de 0 a 11 amostras na entrada;
- Atraso de tempo de 1 amostra na retroalimentação da saída para camada oculta;
- Utilização de *bias* em todas as camadas.

4.2.3 Comparativo entre Modelos

Os modelos $SARIMA(2, 0, 2)(1, 1, 1)_{31}$ definido pela Equação (4.1) e os modelos com redes neurais TDNN e RNN foram utilizados para prever de 1 (5 minutos) até 12 passos à frente (60 minutos) e comparados tomando como base um pacote de 8000 dados (aproximadamente 28 dias) para sintonia/treino e 8000 dados para validação. Os índices de erros de previsão utilizados foram o MAE ($\frac{1}{N} \sum_{k=1}^N (Z_{k+h} - \hat{Z}_{k+h})$) e o MSE ($\frac{1}{N} \sum_{k=1}^N (Z_{k+h} - \hat{Z}_{k+h})^2$), onde h é o horizonte de previsão. Os resultados estão apresentados nas tabelas 5 e 6.

Tabela 5: Comparativo entre Modelos - MSE

Modelo	Horizonte de Previsão											
	1	2	3	4	5	6	7	8	9	10	11	12
SARIMA	1,62	1,62	1,58	1,55	1,60	1,60	1,59	1,71	36,78	44,54	47,50	25,26
TDNN	1,44	6,32	10,35	15,64	20,20	25,60	28,61	30,46	30,39	31,89	31,50	30,46
RNN	1,56	5,31	12,02	15,02	21,35	24,23	27,63	29,40	31,08	37,59	33,88	35,59

Na Figura 43, pode-se visualizar que o modelo SARIMA adaptativo consegue manter o erro baixo até o horizonte $h = 8$ e que a partir deste ponto o modelo linear é incapaz de ajustar os parâmetros para alcançar uma previsão satisfatória. As redes neurais TDNN e RNN apresentam comportamento muito similar e vão gradativamente perdendo a precisão de acordo com o aumento dos horizontes de previsão.

Tabela 6: Comparativo entre Modelos - MAE

Modelo	Horizonte de Previsão											
	1	2	3	4	5	6	7	8	9	10	11	12
SARIMA	0,90	0,91	0,90	0,89	0,89	0,91	0,91	0,96	4,57	5,02	5,22	3,80
TDNN	0,84	1,78	2,35	2,90	3,32	3,73	3,96	4,10	4,12	4,20	4,18	4,10
RNN	0,85	1,65	2,51	2,85	3,41	3,65	3,89	4,02	4,18	4,72	4,50	4,53

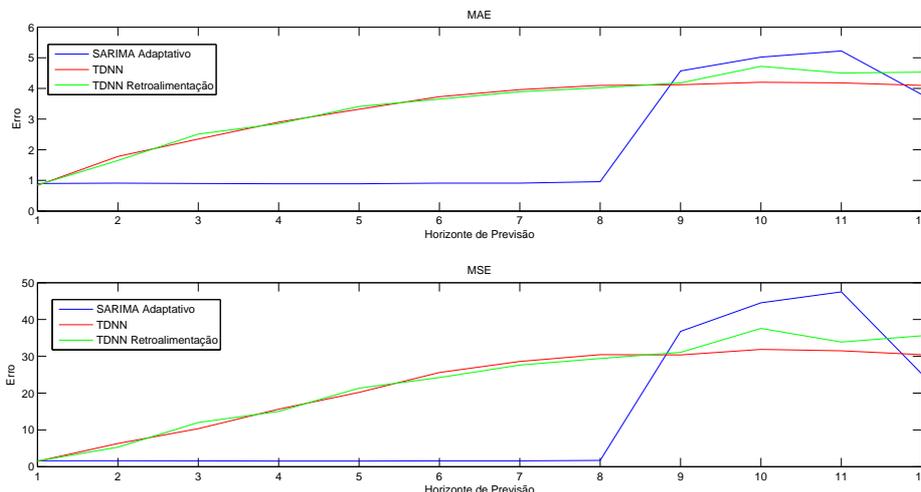


Figura 43: Comparativo entre Modelos - MSE e MAE.

A análise gráfica também é importante para relacionar os valores do MAE e MSE com os erros de previsão nos pontos de mudança de inclinação (vazio e cheio) do nível de líquidos. A previsão para 1 passo à frente pode ser verificada na Figura 44. Para este horizonte, todos os modelos apresentam previsões satisfatórias conseguindo prever todos pontos críticos do processo. Como as redes neurais TDNN e RNN possuem comportamento muito parecidos, apenas inseriu-se nos gráficos a primeira.

Na Figura 45 é possível notar que há uma grande diferença entre as previsões do modelo SARIMA adaptativo e a rede neural TDNN que apresentaram MAEs de 0,96 e 4,10, respectivamente. Essa elevação do erro de 427% além de quantitativamente significativa, pode ser notada qualitativamente no gráfico.

Para o horizonte de 12 passos à frente, mesmo com resultados quantitativos muito próximos, MAEs de 3,80 e 4,10 do SARIMA e do TDNN, respectivamente, a Figura 46 mostra que o primeiro apresenta melhor performance nas mudanças de inclinação e consegue captar melhor as sazonalidades (ciclos) do sinal.

Com base nestas análises foi definido o modelo linear SARIMA adaptativo como o sistema a ser instalado no processo para validação dos resultados.

4.2.4 Validação do Modelo

Após escolha do modelo e do horizonte de previsão (12 passos a frente), foi instalado um computador com um modelo adaptativo na tentativa de se reduzir os erros de previsão, ajustando-se os parâmetros cada vez que novos dados são coletados usando-se uma taxa de amostragem de 5 minutos. Os parâmetros do modelo são calculados através do algoritmo de evolução diferencial usando o erro médio absoluto de previsão para 12 passos a frente como função-custo.

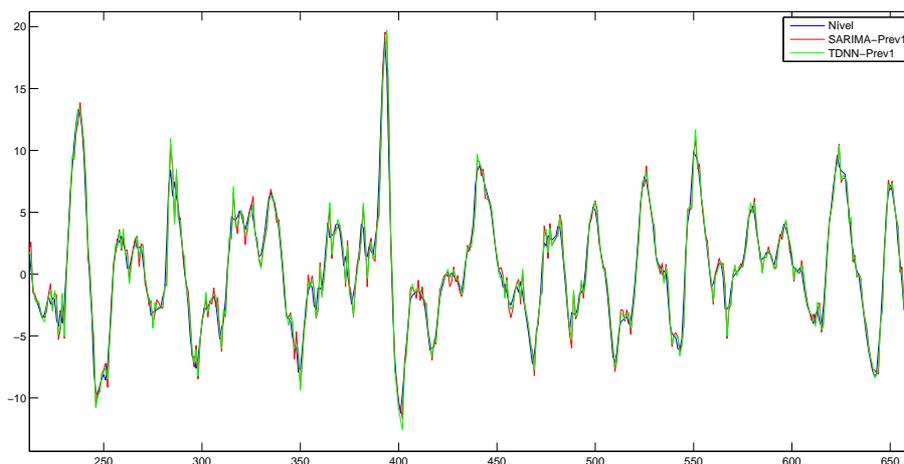


Figura 44: Previsão 1 Passo à Frente - SARIMA x TDNN.

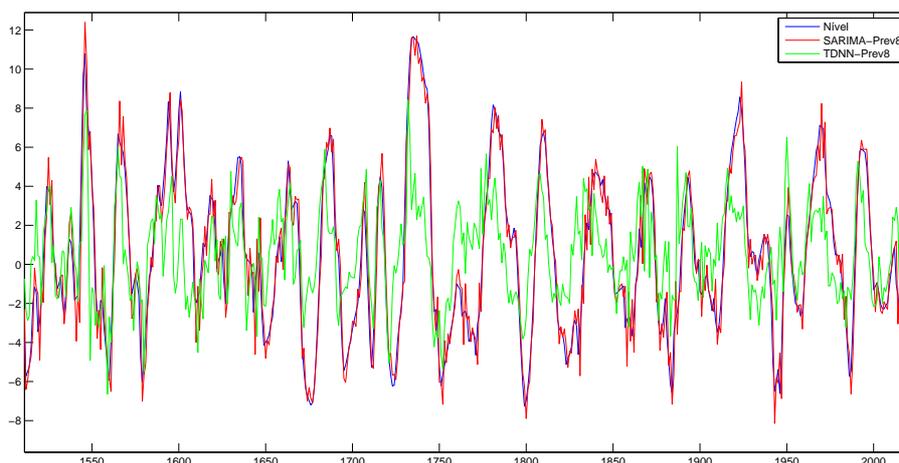


Figura 45: Previsão 8 Passos à Frente - SARIMA x TDNN.

O preditor *on-line* do nível dos líquidos com base na força-eletromotriz (*on-line Multi-step EMF Level Predictor-MELP*) é conectado ao sistema de controle digital do alto-forno 1 através de protocolo OPC (*OLE for Process Control*) que coleta o nível da FEM a cada 5 minutos e devolve a previsão do nível da FEM. A Figura 47 mostra a interface entre a estação MELP e o sistema de controle deste alto-forno.

A escolha do melhor tamanho para o pacote de dados histórico e a quantidade de iterações necessárias para a convergência do algoritmo DE foram obtidos com base na Figura 48. Pode ser observado que depois de 100 iterações do algoritmo DE, todos os pacotes de dados (variando de 2 dias e meio a 40 dias) convergem e que, para pacotes maiores que 20 dias, o algoritmo não mostra reduções na função-custo (MAE). É importante lembrar que os algoritmos genéticos, em geral, representam uma demanda pesada de carga na CPU e que, neste caso, apresenta tempo de processamento de 120 segundos para pacotes de 20 dias com algoritmos DE com 100 iterações. Apesar de parecer grande, estes 2 minutos representam apenas 3.3% de atraso de processamento para um sistema de previsão de uma

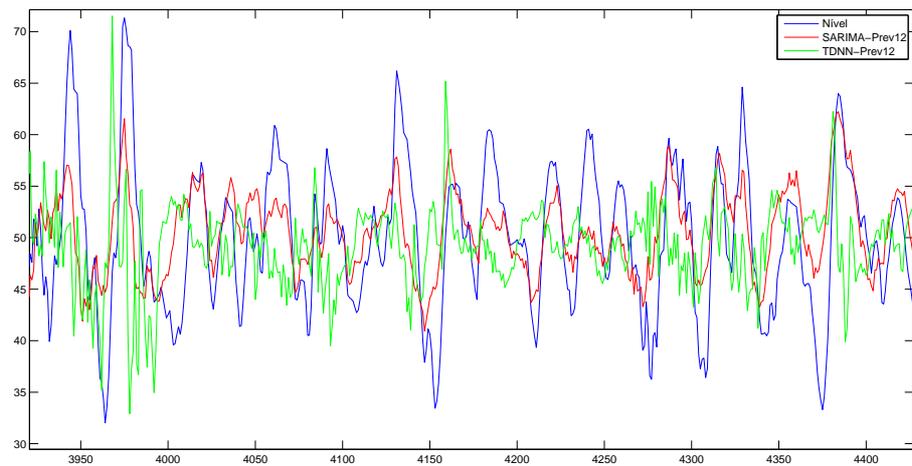


Figura 46: Previsão 12 Passos à Frente - SARIMA x TDNN.

hora à frente (60 minutos).

Com base em ciclos adaptativos de 100 iterações e usando pacotes de dados de 720 minutos, ou 2, 5 dias, o MELP foi simulado para 1 hora à frente apresentando resultados que constata que o modelo mantém o erro de previsão (MAE) sob controle, realizando mudanças nos parâmetros do modelo a cada atualização dos dados. Estes resultados podem ser visualizados na Figura 49.

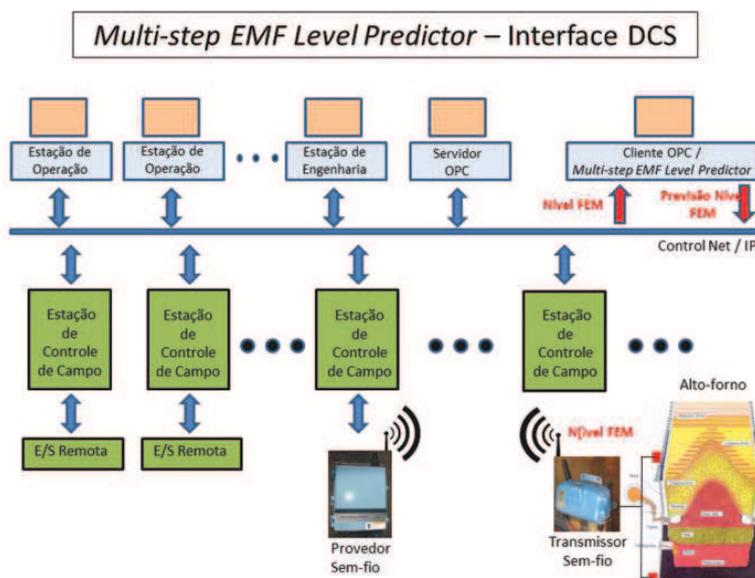


Figura 47: Interface entre MELP e o Sistema de Controle do Alto-forno 1.

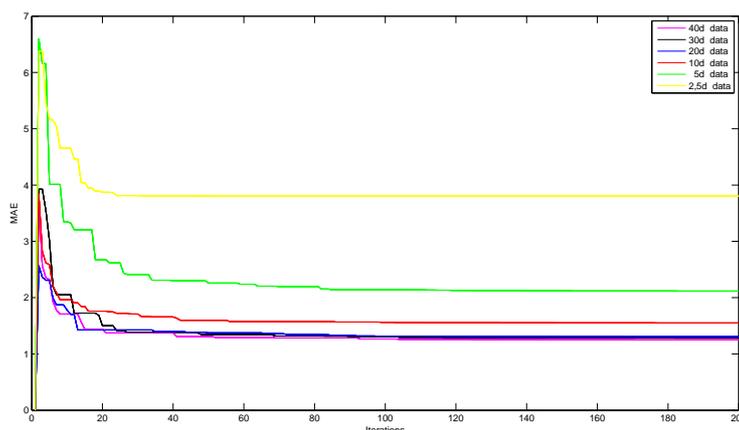


Figura 48: Análise do tamanho de pacotes de dados versus número de iterações.

O sistema de previsão MELP foi então instalado no Alto-forno 1 de uma siderúrgica da Grande Vitória - ES para previsões de 12 passos à frente, 1 hora. Os resultados amostrais obtidos são mostrados na Figura 50. Este modelo adaptativo apresentou 73,38% de precisão (percentual de tempo que o erro de previsão é menor que 10%) e o MAE é igual a 6,98 quando validado em um pacote de 544 pontos. Este resultado indica que os efeitos da não-linearidade podem ser mitigados neste processo utilizando-se este modelo adaptativo para a previsão do nível dos líquidos através da força-eletromotriz na carcaça. Na parte superior da figura, estão mostrados os valores do sinal FEM e das previsões geradas pelo sistema MELP. Na parte inferior, os momentos de abertura e fechamento dos furos-de-gusa sinalizando os instantes em que a extração de líquidos do cadinho está ocorrendo.

É importante notar dois aspectos no comportamento do MELP: primeiro, quando uma anomalia ocorre, como pode ser visualizado na Figura 50 (por volta da amostra 80), quando houve o fechamento emergencial do furo-de-gusa 4 e reabertura quase que imedi-

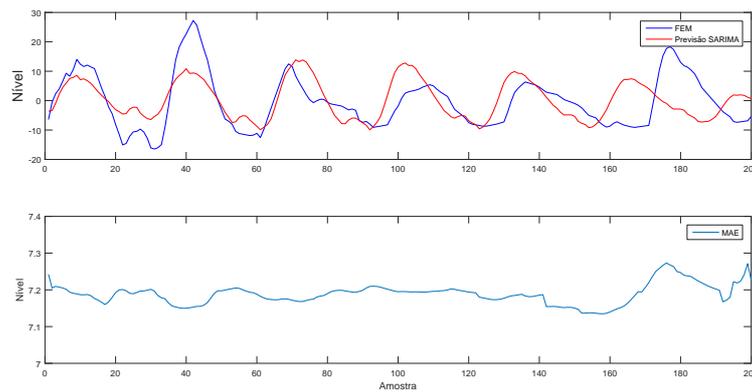


Figura 49: Simulação do MELP para 1 hora à frente com 100 iterações por ciclo.

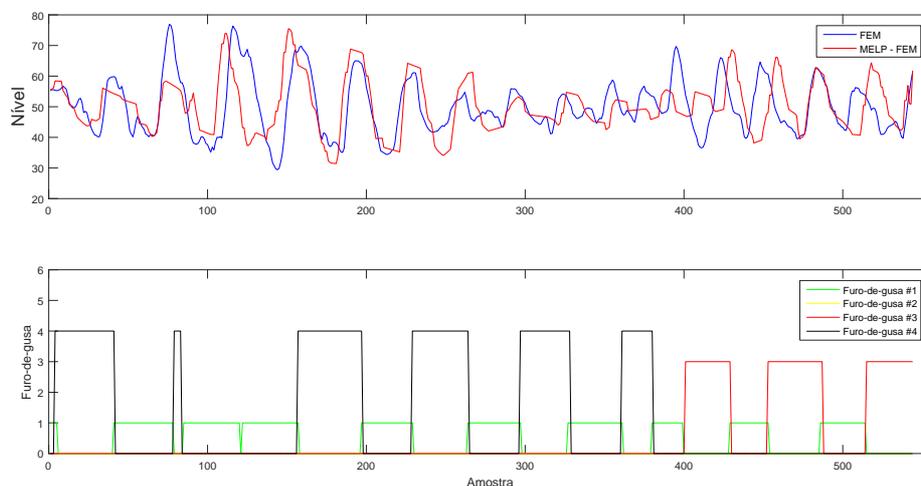


Figura 50: Preditor *On-line* do Nível de Líquidos para 12 passos à frente.

atamente, causando elevação do nível abruptamente, o modelo não foi capaz de se ajustar de imediato, necessitando de alguns ciclos (entre as amostras 80 e 160) para adaptar os parâmetros a esta nova condição. Após este tempo, a previsão de 1 hora à frente da FEM apresentou resultados satisfatórios (entre as amostras 160 e 300).

O segundo aspecto que se deve notar é que, como o modelo MELP é baseado em uma sazonalidade fixa de 31 amostras (ou 155 minutos), ele aumenta muito o erro de previsão quando há mudanças bruscas nos tempos de abertura e fechamento dos furos-de-gusa. Como pode ser visto na Figura 50 (amostras entre 300 e 400), quando houve redução gradual dos tempos das corridas dos furos-de-gusa 1 e 4 e o modelo necessitou de alguns ciclos (depois da amostra 400) para readaptar seus parâmetros e realizar boas previsões.

4.3 Filtragem através do MATLAB

Como o SDCD do alto-forno possui apenas em sua base de dados blocos que representam filtros de primeira ordem, limitando assim a qualidade de filtragem devido à sua reduzida capacidade de corte das frequências próximas às frequências limites do filtro, interferindo assim no sinal filtrado de forma que este possua maior complexidade, o que

pode dificultar o processo de modelagem e, conseqüentemente, elevar os erros de previsão.

Diante desta limitação, optou-se pela filtragem do sinal FEM diretamente no software de previsão. O algoritmo foi desenvolvido de forma que o sistema de filtragem atualiza os dados na mesma frequência de amostragem do sistema de previsão, trabalhando ambos com 300 segundos (5 minutos). Como a maior frequência de interesse será 16 vezes mais baixa que a taxa de amostragem, como será visto a seguir, isto não afeta a informação contida no sinal resultante já que o teorema de Nyquist está sendo respeitado.

Para a realização desta filtragem, foi selecionado um filtro FIR passa-faixa do tipo janela de Kaiser devido sua capacidade de especificação das bandas passantes e oscilações (*ripples*) com precisão. Este filtro foi desenvolvido no software matemático (Matlab) com os seguintes atributos:

- Frequência de corte do passa-baixa em $69,4\mu Hz$ (240 minutos);
- Frequência de passagem do filtro passa-baixas em $83,3\mu Hz$ (200 minutos);
- Frequência de passagem do filtro passa-altas em $0,20mHz$ (83 minutos);
- Frequência de corte do passa-altas em $0,21mHz$ (80 minutos);
- Atenuação das bandas de corte em 60dB;
- Atenuação da banda-passante em 1dB;
- Ordem do filtro de Kaiser 1029;

Após a filtragem, dentro do algoritmo de filtragem, foi removido o atraso de filtragem que corresponde neste caso a 605 lags (3025 minutos ou 50 horas). Uma implicação prática direta da utilização de filtros de ordens elevadas é que, após mudanças ou reinícios no programa, é necessário aguardar este período para geração de *buffer* para eliminação destes atrasos.

O algoritmo de filtragem, que faz parte do algoritmo de previsão, encontra-se na Figura 51.

Este algoritmo de filtragem inicia-se pela abertura do canal de comunicação entre o software (Matlab) e o servidor OPC do DCS. Após isto, realiza-se a atualização do vetor temporal do sinal FEM, fechando-se o link OPC. Na sequência, este sinal é filtrado e devolvido ao DCS através de outra variável (FEM filtrado), após nova abertura do canal OPC. Este ciclo é repetido infinitamente a cada período de 300 segundos (5 minutos).

Na Figura 52 pode-se verificar que o filtro é capaz de eliminar, com precisão, as frequências fora da região de interesse. Em comparação com o sinal da Figura 29, mesmo se tratando de sinais em diferentes períodos de tempo, nota-se que este apresenta energia nas regiões de corte do filtro. Em outras palavras, a complexidade do sinal será maior pois ele possui uma maior quantidade de frequências contidas no sinal a ser modelado e previsto, o que eleva a dificuldade do processo de modelagem e previsão.

Na transformada de Fourier do sinal filtrado é possível verificar a existência de um pico de energia na frequência de $104,27\mu Hz$ (159,84 min). Este pico indica que o período mais frequente entre aberturas/fechamentos dos furos de gusa é de, aproximadamente, 160 minutos. Este valor pode mudar de acordo com as condições operacionais, mas historicamente esta média encontra-se entre 150 e 170 minutos.

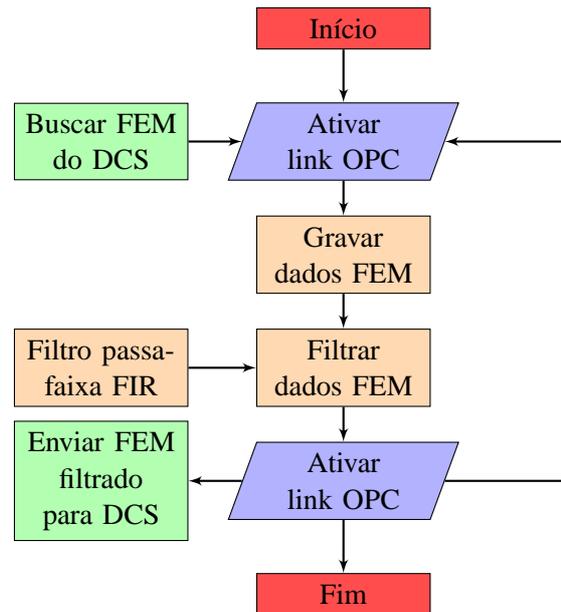


Figura 51: Diagrama da Filtragem do sinal FEM

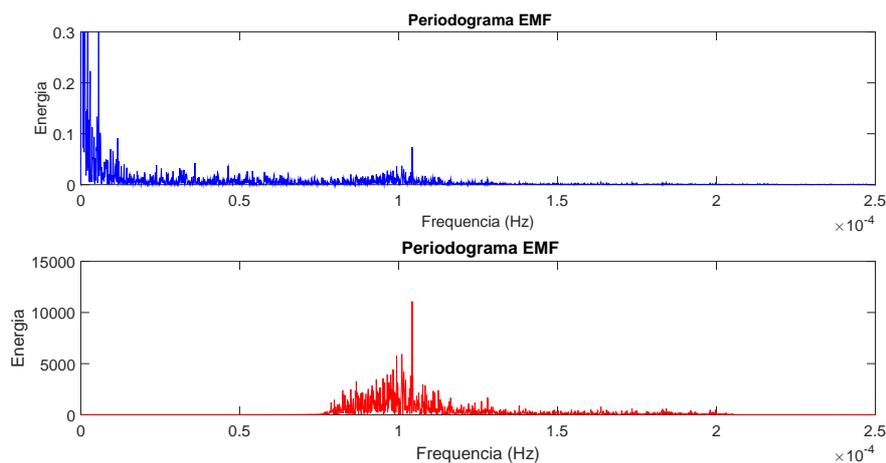


Figura 52: Transformada de Fourier do sinal FEM com filtro FIR.

4.4 Caracterização do Sinal após Nova Filtragem

Como o processo de filtragem afeta as características do sinal significativamente, resolveu-se pela realização de novos testes de estacionariedade e lineariedade através dos testes de KPSS (estacionariedade), Lilliefors e Kruskal-Wallis (lineariedade).

O teste KPSS foi realizado e mostrou que o sinal filtrado pode ser considerado uma série temporal estacionária, como pode ser visto na Tabela 7.

A análise da presença de não-linearidades no sinal FEM filtrado será realizada novamente através dos testes Lilliefors (LILLIEFORS, 1967) e Kruskal-Wallis (KRUSKAL; WALLIS, 1952). Para o último, o pacote de dados foi dividido em oito partes contendo, aproximadamente, 10 dias de dados por pacote. Os resultados destas análises estão contidos na Tabela 8.

Com base nestas análises, pode-se concluir que o sinal, além de ser uma série estacionária, possui componentes de natureza não-linear. Este fato é ratificado através da Figura

Tabela 7: Teste de Estacionariedade - FEM x FEM Filtrado

Sinal	KPSS	Hipótese Nula
FEM	1	$h = 0$, a série temporal é estacionária
FEM Filtrado	0	$h = 1$, a série temporal é raiz unitária

Tabela 8: Teste de Linearidade - FEM Filtrado

Teste	Resultado	Resultados Possíveis
Lilliefors	1	$h = 0$, Dados são normalmente distribuídos $h = 1$, Dados não são normalmente distribuídos
Kruskal-Wallis	0,35	$p \ll 0$, Amostras não possuem mesma distribuição $p \gg 1$, Amostras possuem mesma distribuição

53 que mostra que o sinal apresenta comportamento não-linear nos valores próximos aos limites (cadinho vazio ou cheio).

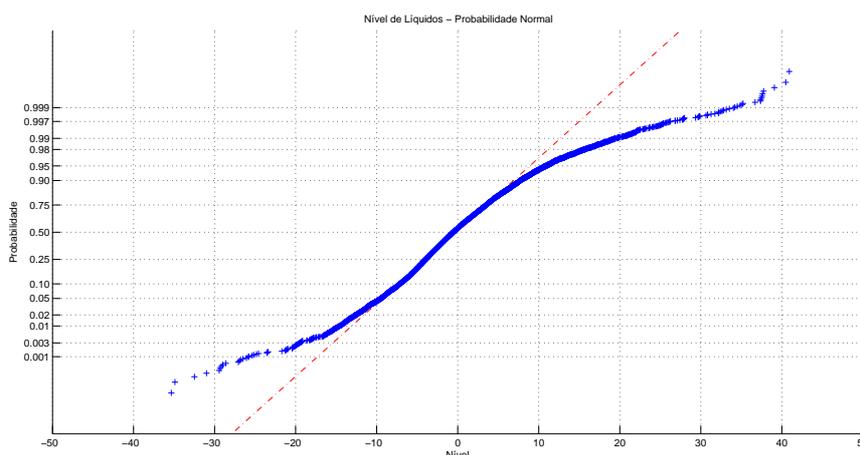


Figura 53: Distribuição normal do FEM filtrado.

As funções de autocorrelação e autocorrelação parcial do novo sinal filtrado, apresentadas na Figura 54, possuem baixo decaimento e oscilações acima dos intervalos de confiança. Os ciclos entre os maiores picos da função de autocorrelação estão entre 31 e 36 atrasos (155 e 180 minutos).

O método de detecção de oscilação em sinais desenvolvido em (KARRA; KARIM, 2009) foi utilizado para identificação das frequências dominantes do sinal com intuito de caracterização da presença de sazonalidades do sinal filtrado já que, pelo conhecimento do processo, espera-se que isto aconteça em virtude do ciclo de aberturas e fechamentos dos furos-de-gusa que, como já citado anteriormente, encontra-se entre 150 e 170 minutos.

O sinal FEM filtrado foi testado para identificação de todas as frequências com potências superiores a 3% da energia total do sinal e regularidade acima de 1. Os resultados encontram-se na Tabela 9.

Este teste identificou a presença da frequência em $104,77\mu Hz$ (159,07 minutos) que carrega 23,69% da energia total do sinal FEM. Esta oscilação pode ser interpretada como uma característica sazonal.

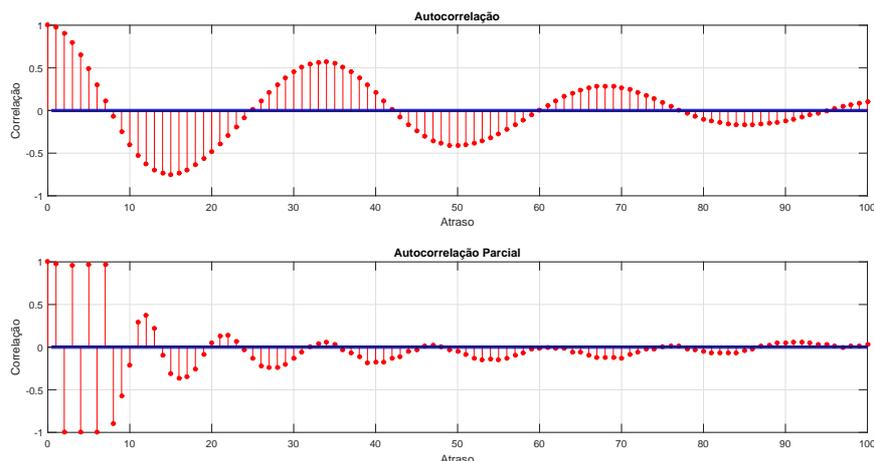


Figura 54: Autocorrelação e autocorrelação parcial do sinal FEM filtrado.

Tabela 9: Detecção de oscilação no FEM filtrado

Frequência	Regularidade	Energia
$78,78\mu Hz$	9,56	5,01%
$100,52\mu Hz$	6101,27	8,40%
$104,77\mu Hz$	4,62	23,69%

Todos estes comportamentos observados durante os testes sugerem a presença de não-linearidades (G.KIRCHGASSNER; WOLTERS, 2007). Também nota-se que o comportamento da curva de autocorrelação e a distribuição de energia da transformada de Fourier indicam múltiplas sazonalidades (KARRA; KARIM, 2009) inerentes ao processo de extração de ferro-gusa e escória do cadinho do alto-forno. Os sinais mais fortes encontrados a cada 32 atrasos (159.84 min) ou frequência de $104,27\mu Hz$ e na distribuição espectral de potência em $104,77\mu Hz$ com base nas análises de oscilações presentes nas Figuras 52 e Tabela 9, respectivamente, ratificam a existência de forte sazonalidade nesta faixa de frequência.

Com a reconfirmação da sazonalidade, o sinal foi novamente diferenciado e submetido a novos testes de linearidade, sendo os resultados apresentados na Tabela 10. Os resultados mostram que o sinal continua não linear após diferenciação mas que possui média e variância constantes, caracterizando-se novamente como um sinal estacionário.

Tabela 10: Teste de linearidade - Sinal filtrado e diferenciado

Teste	Resultado	Resultados Possíveis
Lilliefors	1	$h = 0$, Dados são normalmente distribuídos $h = 1$, Dados não são normalmente distribuídos
Kruskal-Wallis	0,92	$p \ll 0$, Amostras não possuem mesma distribuição $p \gg 1$, Amostras possuem mesma distribuição

Esta não linearidade novamente pode ser visualizada na distribuição da curva normal apresentada na Figura 55. Nota-se que, novamente, existe a presença de não-linearidades nas bordas (nível cheio e vazio do cadinho).

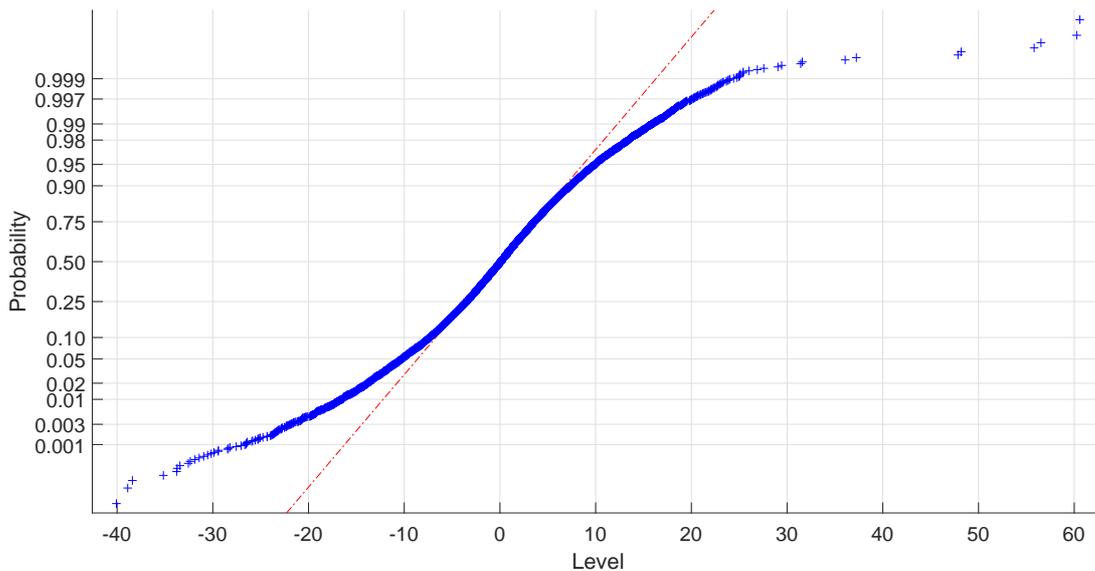


Figura 55: Comportamento do novo sinal filtrado e diferenciado.

4.5 Modelagem e Previsão após Nova Filtragem

Com base nos resultados da Seção 4.4, a modelagem e previsão foi realizada somente considerando os modelos lineares SARIMA (série temporal sazonal autoregressiva e com média móvel) e não-linear TDNN (rede neural com atraso de tempo).

4.5.1 Aproximação Linear

A seleção da ordem dos polinômios $(p, 0, q)$ é realizada com base na análise dos resíduos (NOCEDAL; WRIGHT, 2006) e foi baseada, desta vez, no erro médio quadrático ($MSE(h)$) e no erro médio absoluto ($MAE(h)$) de h passos à frente e comparados para diferentes ordens de modelos ARMA. A ordem do modelo escolhida será aquela que apresentar os menores índices $IMSE(h)$ e $IMAE(h)$, conforme equações abaixo.

$$IMSE(h) = 100 \times \frac{MSE(h)}{MSE_{min}(h)}, \quad IMAE(h) = 100 \times \frac{MAE(h)}{MAE_{min}(h)}$$

onde $MSE_{min}(h)$ e $MAE_{min}(h)$ são, respectivamente, os menores $MSE(h)$ e $MAE(h)$ analisados. O modelo com melhor previsão receberá o índice de 100% de desempenho.

A comparação entre as previsões dos modelos foi executada com base no horizonte de previsão de doze passos à frente (uma hora) e testada até modelos de terceira ordem $(p,0,q)$. Também para esta análise, os parâmetros de cada modelo ARMA foi determinado através do estimador de máxima verossimilhança (ML) com o algoritmo de programação sequencial quadrática (SQP) descrito no Subseção 2.6.3 usando como base o método do gradiente para minimizar o índice $MSE(1)$ do erro residual $re(k+1)$.

Neste teste, um modelo autoregressivo de terceira ordem (ARMA(3,0,0)) apresentou os melhores índices normalizados $IMSE(12)$ e $IMAE(12)$, como apresentado na Tabela 11.

Com base na concentração de energia identificado na transformada de Fourier no Capítulo 4.5 e do algoritmo de detecção de picos no Seção 4.6 que indicam a existência de sazonalidade de 32 lags (160 minutos) neste sinal, classes de modelos sazonais autoregressivos, integrados e de médias móveis (SARIMA(p,d,q)(P,D,Q)₃₂) foram novamente

Tabela 11: IMSE(12) e IMAE(12) para ARMA($p, 0, q$)

Modelo	IMSE (%)	IMAE (%)
(0, 0, 1)	479, 95	261, 90
(0, 0, 2)	398, 76	197, 21
(0, 0, 3)	337, 21	165, 66
(1, 0, 0)	101, 06	103, 50
(1, 0, 1)	100, 94	103, 54
(1, 0, 2)	100, 92	103, 58
(1, 0, 3)	100, 94	103, 66
(2, 0, 0)	174, 40	126, 11
(2, 0, 1)	173, 87	125, 92
(2, 0, 2)	244, 43	146, 86
(2, 0, 3)	217, 03	138, 86
(3, 0, 0)	100, 00	100, 00
(3, 0, 1)	134, 19	113, 01
(3, 0, 2)	154, 69	120, 77
(3, 0, 3)	240, 74	146, 25

testados como uma nova opção de modelos lineares. A Tabela 12 apresenta os índices IMSE(12) e IMAE(12) normalizados para modelos SARIMA($3, 0, 0$)($P, 1, Q$)₃₂ com uma diferenciação sazonal de 32 atrasos de tempo com (p,q) testadas até ordem 3. Os parâmetros são determinados através de estimador ML, com SQP usado para minimizar o índice MSE(1) do erro residual $re(k+1)$ e previsões de 12 passoa à frente.

Tabela 12: IMSE(12) e IMAE(12) para SARIMA($p, 0, q$)($0, 1, 0$)₃₂

Modelo	IMSE (%)	IMAE (%)
(0, 0, 1)(0, 1, 0) ₃₂	121, 76	106, 29
(0, 0, 2)(0, 1, 0) ₃₂	121, 76	106, 29
(0, 0, 3)(0, 1, 0) ₃₂	121, 76	106, 29
(1, 0, 0)(0, 1, 0) ₃₂	250, 70	151, 98
(1, 0, 1)(0, 1, 0) ₃₂	238, 75	148, 11
(1, 0, 2)(0, 1, 0) ₃₂	235, 46	146, 58
(1, 0, 3)(0, 1, 0) ₃₂	214, 64	140, 57
(2, 0, 0)(0, 1, 0) ₃₂	104, 29	102, 35
(2, 0, 1)(0, 1, 0) ₃₂	103, 88	102, 11
(2, 0, 2)(0, 1, 0) ₃₂	105, 75	102, 97
(2, 0, 3)(0, 1, 0) ₃₂	106, 49	103, 28
(3, 0, 0)(0, 1, 0)₃₂	100, 00	100, 00
(3, 0, 1)(0, 1, 0) ₃₂	100, 46	100, 26
(3, 0, 2)(0, 1, 0) ₃₂	103, 25	101, 65
(3, 0, 3)(0, 1, 0) ₃₂	120, 14	109, 77

Após a seleção do modelo SARIMA($3, 0, 0$)($0, 1, 0$)₃₂ como o de melhor desempenho, testou-se desta vez modelos com variações de (P,Q) até ordem 3, ou seja, modelos SARIMA($3, 0, 0$)($P, 1, Q$)₃₂, resultando na Tabela 13. O modelo que apresentou melhores resultados em ambos os critérios (IMSE(12) e IMAE(12)) foi o SARIMA($3, 0, 0$)($1, 1, 2$)₃₂.

Tabela 13: IMSE(12) e IMAE(12) para SARIMA(3, 0, 0)(P, 1, Q)₃₂

Model	IMSE (%)	IMAE(%)
(3, 0, 0)(0, 1, 0) ₃₂	115, 83	107, 87
(3, 0, 0)(0, 1, 1) ₃₂	142, 35	117, 95
(3, 0, 0)(0, 1, 2) ₃₂	100, 86	101, 06
(3, 0, 0)(0, 1, 3) ₃₂	100, 59	100, 03
(3, 0, 0)(0, 1, 0) ₃₂	142, 35	117, 95
(3, 0, 0)(1, 1, 1) ₃₂	105, 07	102, 47
(3, 0, 0)(1, 1, 2)₃₂	100, 00	100, 00
(3, 0, 0)(1, 1, 3) ₃₂	139, 43	117, 12
(3, 0, 0)(2, 1, 0) ₃₂	151, 13	121, 98
(3, 0, 0)(2, 1, 1) ₃₂	101, 70	100, 45
(3, 0, 0)(2, 1, 2) ₃₂	151, 25	122, 07
(3, 0, 0)(2, 1, 3) ₃₂	155, 79	124, 47
(3, 0, 0)(3, 1, 0) ₃₂	270, 61	162, 28
(3, 0, 0)(3, 1, 1) ₃₂	181, 69	133, 90
(3, 0, 0)(3, 1, 2) ₃₂	204, 80	141, 03
(3, 0, 0)(3, 1, 3) ₃₂	234, 87	152, 08

4.5.2 Estimação de Parâmetros do Modelo SARIMA com Algoritmo Genético

O próximo passo na análise de um modelo linear é realizar a estimação do modelo SARIMA escolhido através de algoritmos genético, DE no caso, e comparar seu desempenho com o desempenho do mesmo modelo com parâmetros estimados através de SQP.

Neste caso, o algoritmo DE apresentou vantagens em relação ao método SQP para estimação de parâmetros como mostrado na Figura 56, onde o desempenho de previsão dos modelos e seus respectivos estimadores são comparados usando-se o índice MSE(h) para modelos ARMA(3, 0, 0), denominados ARMA-SQP-1, ARMA-DE-1 e ARMA-DE-h, onde seus parâmetros foram estimados minimizando-se a função MSE(1) do SQP e DE para o horizonte de previsão de 1 hora e também a função-custo minimizada no horizonte de previsão desejado (h).

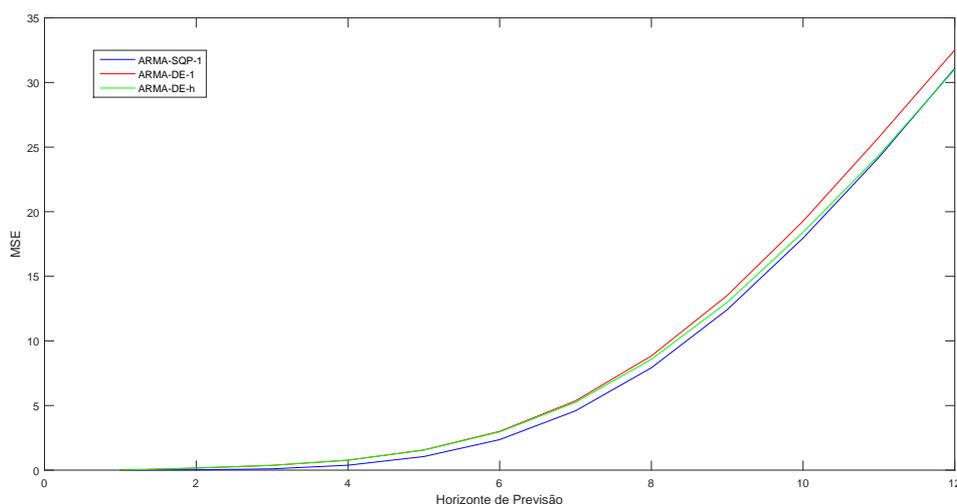


Figura 56: Modelos ARMA com Parâmetros Estimados por algoritmos SQP e DE.

A Figura 57 mostra a mesma comparação mas, neste caso, considera o modelo SARIMA escolhido na Subseção 4.5.1 SARIMA(3, 0, 0)(1, 1, 2)₃₂ comparando os desempenhos entre o SARIMA-SQP-1, SARIMA-DE-1 e SARIMA-DE-h.

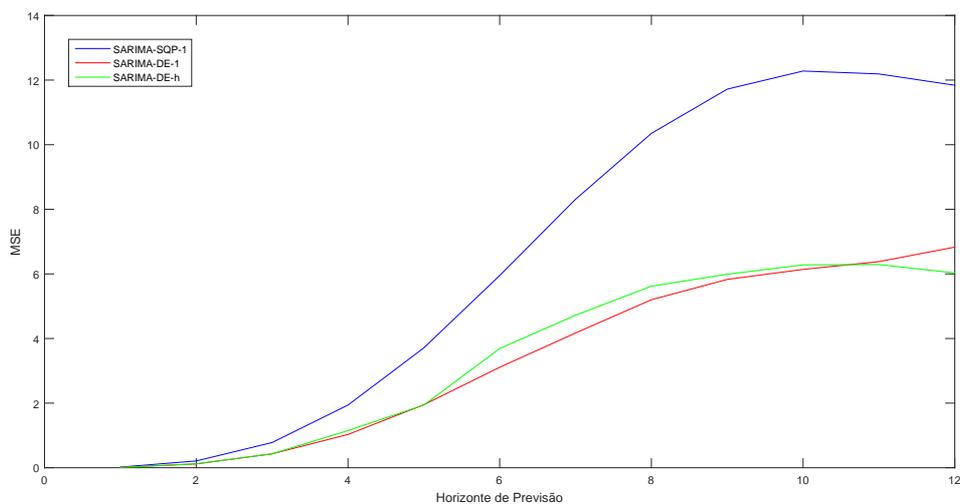


Figura 57: Modelos SARIMA com Parâmetros Estimados por algoritmos SQP e DE.

As Figuras 56 e 57 mostram que o algoritmo DE é capaz de obter significativas melhorias na precisão das previsões para ambos os modelos e que a identificação dos parâmetros dos modelos diretamente nos horizontes de previsão não apresenta ganhos de precisão de desempenho, portanto, não se justifica. Além disso, nota-se que os desempenhos de previsão dos modelos SARIMA são melhores que aqueles apresentados pelos modelos ARMA em todos os casos.

4.5.3 Modelos com Redes Neurais

Para utilização nesta nova FEM filtrada, foram testadas algumas redes neurais com atraso de tempo (TDNNs) com mudanças na quantidade de neurônios e atrasos de tempo com valores de 10 até 70 para avaliação da melhor configuração que apresente o melhor desempenho de previsão para doze passos à frente (60 minutos). Os resultados encontram-se na Tabela 14.

Tabela 14: Modelos TDNN (neurônios,atrasos)

Modelo	IMSE (%)	IMAE(%)
TDNN(10, 10)	2273,33	446,99
TDNN(10, 30)	326,67	175,19
TDNN(10, 50)	220,00	151,13
TDNN(10, 70)	100,00	100,00
TDNN(30, 10)	2646,67	577,82
TDNN(30, 30)	340,00	180,08
TDNN(30, 50)	2000,00	140,23
TDNN(30, 70)	106,67	106,39
TDNN(50, 10)	2700,00	583,08
TDNN(50, 30)	380,00	199,62
TDNN(50, 50)	226,67	153,76
TDNN(50, 70)	120,00	112,78
TDNN(70, 10)	2750,67	184,89
TDNN(70, 30)	400,00	206,77
TDNN(70, 50)	926,67	206,77
TDNN(70, 70)	146,67	128,20

Na Tabela 14 nota-se que o melhor desempenho foi apresentado pela TDNN(10, 70). Assim, foram testadas redes com até 300 atrasos para avaliação dos ganhos de precisão na previsão; Contudo, de acordo com a Figura 58, não há melhoria que justifique atrasos de rede superiores a 70. Dessa forma, a TDNN(10, 70) foi escolhida.

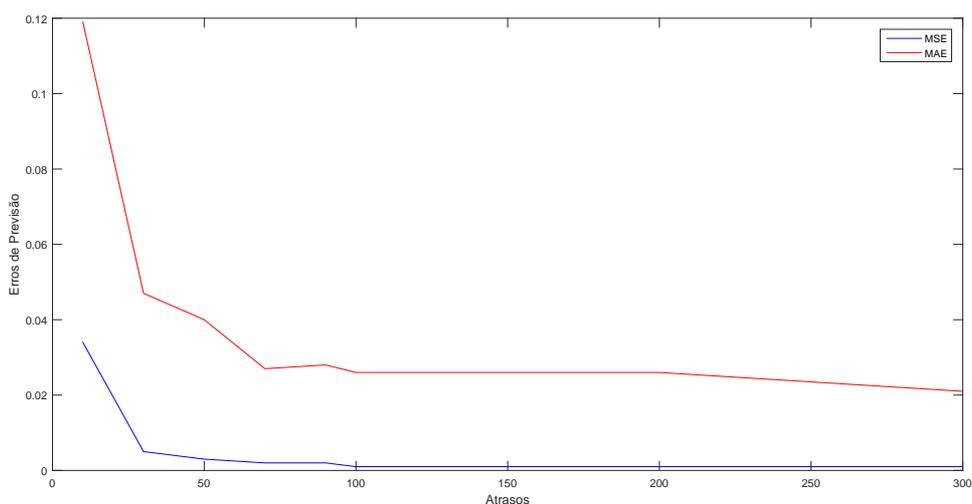


Figura 58: TDNN - Atrasos x Erros de Previsão.

Esta rede TDNN possui as seguintes características:

- 02 camadas, sendo 01 oculta;

- 10 neurônios na primeira camada e 01 neurônio na última camada;
- Funções de transferência logarítmica sigmóide (Logsig) e linear (purelin) na primeira e última camadas, respectivamente;
- Entrada com atrasos de tempo entre 0 – 70 nas amostras;
- Uso de constantes (*bias*) em todas as camadas;
- Uso de metodologia de realimentação de Levenberg-Marquardt como função de treinamento.

4.5.4 Comparação entre os Modelos

O melhor modelo SARIMA, com parâmetros estimados pelo algoritmo DE e o melhor modelo TDNN escolhido pela análise de previsão de doze passos à frente (60 minutos) foram comparados utilizando-se um pacote de dados de validação de 5760 pontos. É importante ressaltar que este pacote de dados não pertence a nenhum pacote de dados presente durante estimação de parâmetros (SARIMA) nem no pacote de treino da rede TDNN. Para comparação, o MSE ($\frac{1}{N} \sum_{k=1}^N (Z_{k+h} - \hat{Z}_{k+h})^2$) e MAE ($\frac{1}{N} \sum_{k=1}^N (Z_{k+h} - \hat{Z}_{k+h})$) de h passos à frente são usados como métricas, onde h é o horizonte de previsão, e N é o número de pontos de validação.

Como mostrado nas Figuras 59 e 60, o modelo TDNN apresentou um desempenho muito superior em comparação ao modelo linear e manteve os erros de previsão baixos até doze horizontes à frente (MSE e MAE aproximadamente 0,01 e 0,27 para 12 passos à frente, respectivamente). O modelo linear apresentou significativa perda de precisão com o aumento do horizonte de previsão enquanto que a RNA consegue manter os erros sob controle.

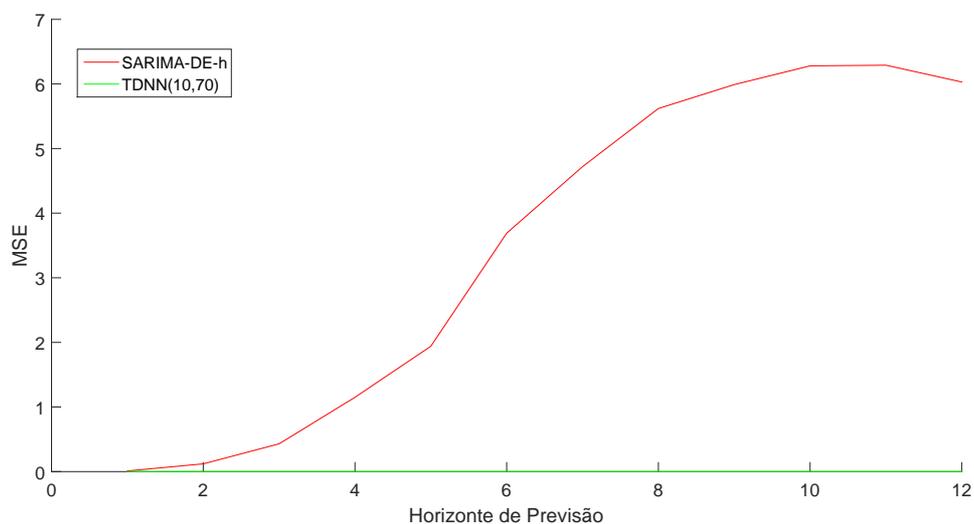


Figura 59: Comparação entre Modelos de Previsão - MSE.

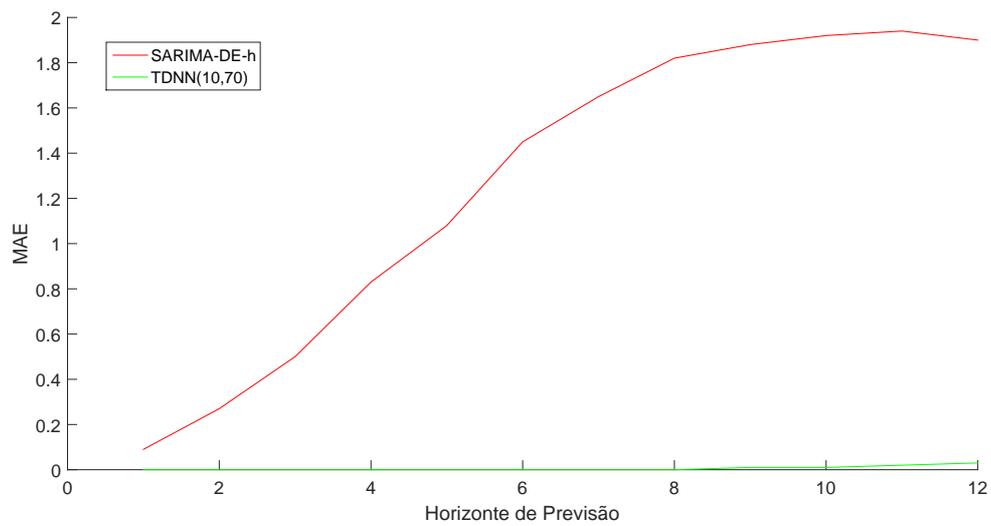


Figura 60: Comparação entre Modelos de Previsão - MAE.

As Figuras 61 e 62 incluem uma amostra de 300 pontos (25 horas) dos dados de validação que mostra a análise temporal e os erros para previsões de 12 passos à frente, respectivamente. Nota-se que o sinal EMF (linha preta) e a previsão da rede TDNN são muito próximas, dificultando a visualização.

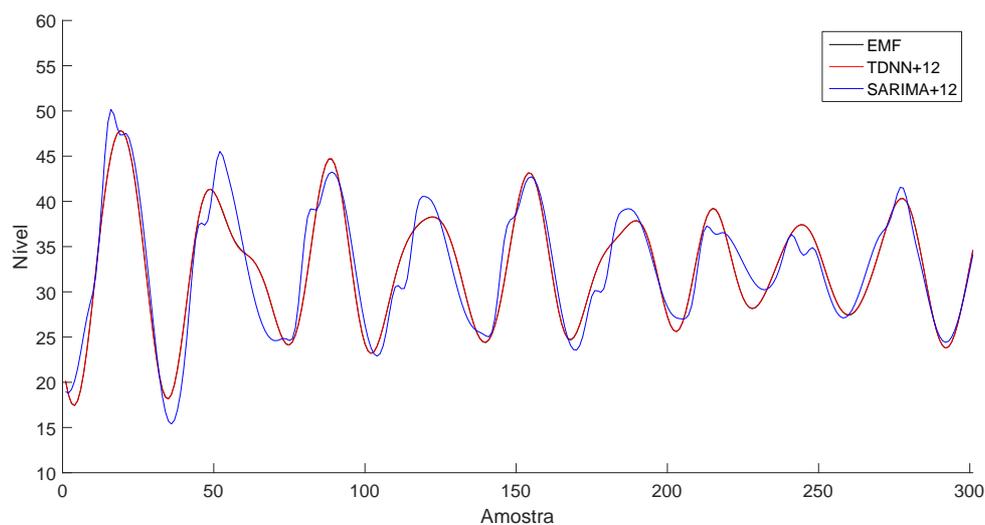


Figura 61: Comparação entre Modelos de Previsão - Análise Temporal para 12 Passos à Frente.

Na análise temporal das Figuras 61 e 62 é interessante relacionar também, os valores de MAE e MSE, além dos erros de previsão, com os pontos de mudança de inclinação da curva de nível de líquidos (vazio e cheio), mostrando as forças e fraquezas dos modelos.

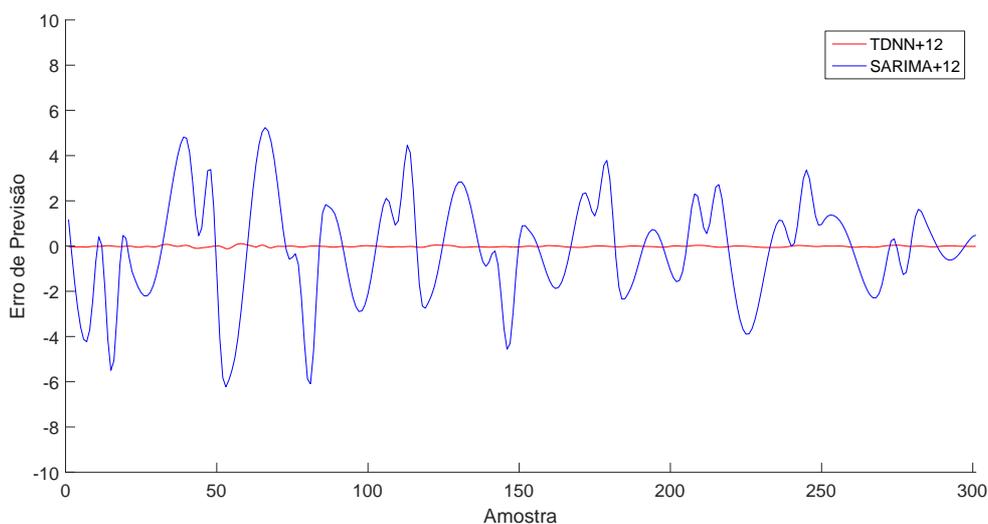


Figura 62: Comparação entre Modelos de Previsão - Erros para 12 Passos à Frente.

Com base nestas análises, o modelo TDNN foi definido como o sistema a ser instalado no processo em tempo real para validação dos resultados. Uma possível explicação para o baixo desempenho apresentado pelo modelo SARIMA é que modelos lineares não conseguem capturar essas relações. Esta característica pode ser, claramente, observada em previsões para horizontes maiores.

4.6 Validação do Modelo em Tempo Real

A validação dos resultados anteriores acontece com a implantação de um computador executando o modelo de rede neural coletando os dados do processo produtivo com taxa de amostragem de 300 segundos (5 minutos). Desta vez, a rede TDNN foi treinada *offline* considerando como função-custo, o erro absoluto médio de previsão para doze passos à frente como discutido na Seção 4.4.4 e 4.7.4.

O preditor de nível dos líquidos com base na força-eletromotriz (*Multi-step EMF Level Predictor- MELP*) foi conectado ao sistema de controle digital da mesma forma do teste de validação anterior mostrado na Figura 47.

Contudo, a Figura 63 mostra que o algoritmo de previsão MELP opera em paralelo com o algoritmo de filtragem MELP mostrado anteriormente na Figura 51. o Algoritmo de previsão MELP consiste na obtenção de um sinal FEM filtrado do sistema de controle digital do alto-forno e com seu histórico gravado em um vetor FEM registrado a cada intervalo definido (300 segundos neste caso). Após isto, o sinal FEM alimenta a rede TDNN, gerando as previsões. Estas previsões produzidas no processo são entregues ao DCS do alto-forno como uma nova variável.

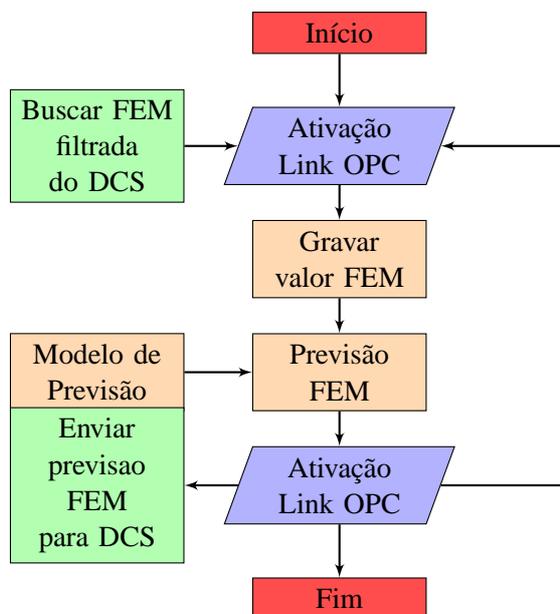


Figura 63: Diagrama de Previsão do Sinal EMF

O sistema MELP para previsão de uma hora à frente do sinal EMF em ação é mostrado na Figura 64 com validação utilizando-se um pacote de 1728 pontos (6 dias). Este modelo apresentou MSE e MAE iguais a 0,43 e 0,63, respectivamente. A mesma figura mostra o desempenho do modelo em relação à previsão dos picos e vales. Neste caso, o modelo apresentou MSE e MAE iguais a 0,42 e 0,62, respectivamente, um erro médio absoluto de tempo de 0 atrasos com erros máximos de 2 atrasos (10 minutos) em comparação com os picos e vales reais.

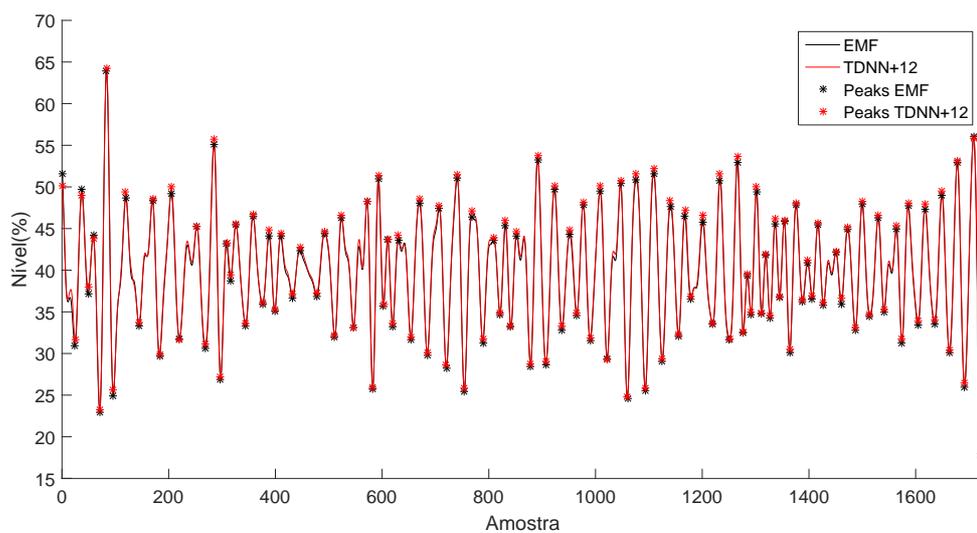


Figura 64: Previsão on-line do MELP para 12 passos à frente.

A Figura 65 mostra os erros de previsão para o mesmo pacote de dados. O modelo TDNN pode manter o erro médio absoluto sob controle (menor que 2%), com erro médio de -0.69% .

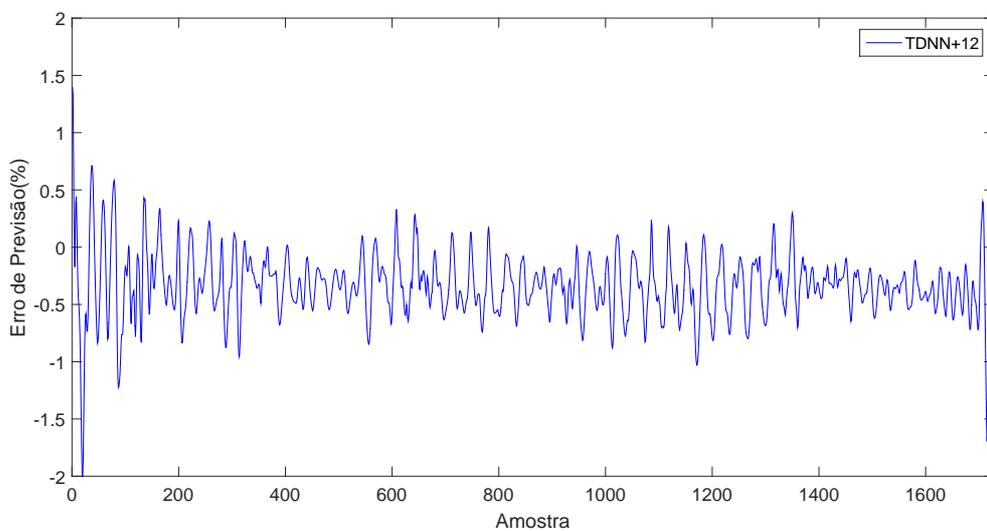


Figura 65: Erros de previsão on-line do MELP para 12 passos à frente.

5 CONCLUSÃO

O desenvolvimento de um modelo com base em séries temporais para a previsão dos níveis dos líquidos no interior de um alto-forno industrial real, para períodos de até 1 hora, utilizando-se da força-eletromotriz gerada na carcaça como fonte de informação resultou em um procedimento de fácil execução e que pode ser replicado para outras plantas.

Com base nos dados obtidos das simulações e observações, tem-se algumas conclusões a respeito da força-eletromotriz:

- O estudo mostrou que o sinal sem tratamento era não-estacionário e não-linear. Após filtragem descrita no Capítulo 4, este novo sinal se tornou estacionário mantendo ainda o comportamento não-linear.
- O sinal da FEM tem um comportamento sazonal e não-linear, que são relacionados ao processo de extração do ferro-gusa e escória do cadinho do alto-forno.

Em virtude da presença de não-linearidade trabalhou-se com duas metodologias distintas para abordar o problema: uma abordagem com modelos lineares (ARMA e SARIMA) com parâmetros estimados através de algoritmos genéticos e outra com redes neurais artificiais (TDNN) na tentativa de capturar estas não-linearidades.

Estes dois métodos foram comparados quantitativa e qualitativamente quanto as suas capacidades de previsão de até 1 hora à frente. A filtragem do sinal da FEM através de um filtro FIR passa-faixa tipo janela de Kaiser permitiu mostrar que a rede neural com atraso de tempo apresentou melhor desempenho na previsão do sinal até 1 hora à frente que os modelos lineares ARMA e SARIMA. Por outro lado, quando a filtragem da FEM é realizada pelo filtro de primeira ordem, não houve condições da rede neural representar todos os sinais contidos na FEM, apresentando previsões piores que os modelos lineares.

O sistema de previsão usando o filtro FIR passa-faixa tipo janela de Kaiser e a rede neural TDNN foi implantado numa planta produtiva real gerando previsões de forma sistemática e contínua utilizando-se da capacidade de adaptação do modelo quanto às mudanças no processo. O modelo TDNN proposto provou ser uma ferramenta promissora para a previsão do nível de líquidos no cadinho do alto-forno.

A informação de uma hora a frente pode ajudar operadores e engenheiros durante o controle e otimização da produção dos altos-fornos, por exemplo, quando enfrentar problemas de retenção de líquidos, durante análises de balanço de massa e energia, no auxílio quando houver necessidade de mudanças nas taxas de produção e para identificação de problemas de manutenção. Em outras palavras, esta informação fornece segurança adicional e ganhos financeiros para o processo do alto-forno.

5.1 Recomendações para Trabalhos Futuros

A metodologia desenvolvida pode ser melhorada com foco na expansão da capacidade de previsão e incremento da precisão nos pontos críticos do processo (níveis cheio e vazio).

Destas, é importante citar algumas possibilidades:

- Caracterização do tipo de não-linearidade com foco na utilização de modelos não lineares de séries temporais como, por exemplo, modelos ARCH, GARCH, etc;
- Avaliação de modelos de séries temporais com memória longa;
- Avaliação da série temporal através de redes wavelets visando melhoria da filtragem e, conseqüentemente, dos resultados da previsão;
- Estudo para implantação de um modelo evolutivo de redes neurais, com objetivo de se selecionar a melhor estrutura de rede neural;
- Avaliação de implantação de modelos mistos tais como: neural-fuzzy, neural-SARIMA dentre outros com o intuito de soma das melhores características de cada sistema;
- Utilização de técnicas de classificação para indicar o momento de abertura e/ou fechamento dos furos-de-gusa, podendo ser testadas redes neurais, SVM, KNN, árvores de decisão, etc.

REFERÊNCIAS

- ABO-HAMMOUR, Z.; ALSMADI, O.; A.M.AL-SMADI; ZAQOUT, M.; SARAIHEH, M. ARMA model order and parameter estimation using genetic algorithms. **Mathematical and Computer Modelling of Dynamical Systems**, [S.l.], n.2, p.201–221, 2012.
- AGUILAR, J.; TURIAS, I.; M.J.JIMENEZ-COME. Hybrid approaches based on SARIMA and artificial neural networks for inspection time series forecasting. **Transportation Research Part E**, [S.l.], n.67, p.1–13, 2014.
- AGUIRRE, L. **Introdução à Identificação de Sistemas: técnicas lineares e não-lineares aplicadas a sistemas reais**. 2nd.ed. Belo Horizonte, Brasil: Editora UFMG, 2007.
- ALTER, M.; BRUNNER, J.; HOMES, D. Continuous Monitoring of Liquid Level and Thermal State in the Hearth Based on Measurement of EMF on the Blast Furnace Shell. **Iron & Steel Technology**, [S.l.], p.43–50, 2013.
- BOX, G.; JENKINS, G.; REINSEL, G. **Time Series Analysis: forecasting and control**. 4nd.ed. USA: Cliffs:Prentice Hall, 1994.
- BRAMMING, M.; HALLIN, M. **Development of new techniques for blast furnace process control**. S.S.A.B. Lulea Works.
- BRANNBACKA, J.; SAXEN, H. Modeling the liquid levels in the blast furnace hearth. **ISIJ International**, [S.l.], n.41, p.1131–1138, 2001.
- BRANNBACKA, J.; SAXEN, H. Novel Model for Estimation of liquid Levels in the Blast Furnace Hearth. **Chemical Engineering Science**, [S.l.], n.59, p.3423–3432, 2004.
- BRANNBACKA, J.; SAXEN, H. Dynamic model of liquid level in the blast furnace hearth. **Scandinavian Journal of Metallurgy**, [S.l.], p.116–121, 2005.
- CHEN, B.; LEE, B.; PENG, S. Maximum Likelihood Parameter Estimation of F-ARIMA Processes Using the Genetic Algorithm in the Frequency Domain. **IEEE Transactions on Signal Processing**, [S.l.], n.50, p.2208–2220, 2002.
- CORDUAS, M. Nonlinearity Tests in Time Series Analysis. **Italian Statistics Society**, [S.l.], n.03, p.291–313, 1994.
- DEPIZZOL, C. F. **Contribuições ao Problema de detecção Automática de oscilações em Malhas de Controle**. Vitória, Brasil: Universidade Federal do Espírito Santo, 2011.

DONATE, J. P.; LI, X. Time series forecasting by evolving artificial neural networks with genetic algorithms, differential evolution and estimation of distribution algorithm. **Neural Computational and Applications**, [S.l.], n.22, p.11–20, 2013.

DOROFEEV, V.; NOVOKHATSKII, A. Origin of difference in electrical potentials on blast furnace shell. **Steel in teh U.R.S.S.**, [S.l.], n.01, 1984.

DUBOVIK, V. Automatic monitoring of the state of the blast furnace wall. **Izv. V.U.Z. Chernaya Metall**, [S.l.], 1982.

ELSAYED, S. M.; SARKER, H. A.; ESSAM, D. L. An Improved Self-Adaptive Differential Evolution Algorithm for Problems. **IEEE Transactions on Industrial Informatics**, [S.l.], n.09, 2013.

FARAHAT, J. V.; MCDONALD, J. B.; NELSON, R. D. Time Series Prediction with Genetic-algorithm Designed Neural Networks: an empirical comparison with modern statistical models. **Computational Intelligence**, [S.l.], n.3, 1999.

FARAHAT, M. A.; TALAAT, M. Short-Term Load Forecasting Using Curve Fitting Prediction Optimized by Genetic Algorithms. **International Journal of Energy Enginerring**, [S.l.], p.23–28, 2012.

FARUK, D. O. A hybrid neural network and ARIMA model for water quality time series prediction. **Engineering Applications of Artificial Intelligence**, [S.l.], n.23, p.586–594, 2010.

FRANSSEN, C. Application of different measuring techniques for the evaluation of hearth permeability. **Proc. 3rd International Conference on Science and Technology of Ironmaking**, [S.l.], p.471–476, 2003.

FUJIHARA, F. **Aspectos Tecnológicos Relativos à preparação de Carga e Operação de Alto Forno**. Associação Brasileira de Metalurgia e Materiais, ABM.

GEERDES, M.; VLIET, C.; TOXOPEUS, H.; F.T.P.MEDEIROS. **Prática Modernas para Operação de Altos-fornos**. xxx.ed. São Paulo, Brasil: Editora Universidade Deft, 2007.

G.KIRCHGASSNER; WOLTERS, J. **Introduction to Modern Time Series Analysis**. Heidelberg, Germany: Springer-Verlag Berlin, 2007.

GOLSHAN, M.; SAMET, H. Updating stochastic models of arc furnace reactive power by genetic algorithm. **14th International Conference on Harmonics and Quality of Power (ICHQP)**, [S.l.], p.1–9, 2010.

GOMES, F. **Ferramenta Preditiva para o Nível de Líquidos em Cadinhos de Altos-fornos**. Vitória, Brasil: Universidade Federal do Espírito Santo, 2010.

GOMES, F. S. V.; COCO, K. F.; SALLES, J. L. F. Multistep Forecasting Models of the Liquid Level in a Blast Furnace Hearth. **IEEE Transactions on Automation Science and Engineering**, [S.l.], 2016.

GOMES, F.; SALLES, J. A New Prediction Model for Liquid Level in Blast Furnaces Based on Time Series Analysis. **9th IEEE International Conference on Control and Automation (ICCA)**, [S.l.], p.772–777, 2011.

HAMEDANI, M. E.; SAMET, H. Updating Stochastic Models of Arc Furnace Reactive Power by Genetic Algorithm. **14th International Conference on Harmonics and Quality of Power**, [S.l.], 2010.

HAMILTON, J. **Time Series Analysis**. 1st.ed. New Jersey,USA: Princeton University Press, 1994.

HAYKIN, S. **Neural Networks: a comprehensive foundation**. 2nd.ed. Ontario, Canada: Prentice Hall, 1998.

HU, Y. H.; HWANG, J. **Handbook of neural network signal processing**. USA: CRC Press, 2002.

INGLE, V. K. **Digital Signal Processing Using Matlab - Third Edition**. 1.ed. Chichester, United Kingdom: John Wiley and Sons, 2012.

JHA, G. K.; SINHA, K. Time-delay neural networks for time series prediction: an application to the monthly wholesale price of oilseeds in india. **Neural Computational and Applications**, [S.l.], n.24, p.563–571, 2014.

KAISER, J. F.; SCHAFFER, R. W. On the Use of the Io-Sinh Window for Spectrum Analysis. **IEEE Transactions on Acoustics, Speech and Signal Processing**, [S.l.], n.1, p.105–107, 1980.

KARRA, S.; KARIM, M. Comprehensive methodology for detection and diagnosis of oscillatory control loops. **Control Engineering Practice**, [S.l.], n.17, p.939–956, 2009.

KHASHEI, M.; BIJARI, M. An Artificial neural network (p,d,q) model for timeseries forecasting. **Expert Systems with Applications**, [S.l.], n.37, p.479–489, 2010.

KHASHEI, M.; BIJARI, M. A New Hybrid Methodology for Nonlinear Time Series Forecasting. **Modelling and Simulation in Engineering**, [S.l.], 2011.

KOLMOGOROV, A. Sulla determinazione empirica di una legge di distribuzione. **G. Ist. Ital**, [S.l.], n.4, p.83–91, 1933.

KRUSKAL, W.; WALLIS, W. Use of ranks in one-criterion variance analysis. **Journal of the American Statistical Association**, [S.l.], n.47, p.583–621, 1952.

KUO, B. **Automatic Control Systems**. 7nd.ed. New Jersey,USA: Prentice Hall, 1995.

KWIATKOWSKI, D.; PHILLIPS, P. C. B.; SCHMIDT, P.; SHIN, Y. Testing the null hypothesis of stationarity against the alternative of a unit root. **Journal of Econometrics**, [S.l.], n.54, p.159–178, 1982.

LEBED, P.; SAFINA, L.; DEMIDENKO, V.; SEMENKO, A.; POCHEKAILO, I. Evaluating work of blast furnace hearth from the nature of EMF variation. **Steel in the U.R.S.S.**, [S.l.], n.20, 1990.

LEICESTER, H. Towards Seamless Computing and Metacomputing in Java. **The Journal of Chemical Education**, [S.l.], n.28, p.581–583, 1951.

- LILLIEFORS, H. W. On the Kolmogorov-Smirnov test for normality with mean and variance unknown. **Journal of the American Statistical Association**, [S.l.], n.62, p.399–402, 1967.
- LIU, H.; TIAN, H.; LI, Y. Comparison of two new ARIMA-ANN and ARIMA-Kalman hybrid methods for wind speed prediction. **Applied Energy**, [S.l.], n.98, p.415–424, 2012.
- LUUKKONEN, R.; SAIKKONEN, P.; TERASVIRTA, T. Testing Linearity in Univariate Time Series Models. **Scandinavian Journal of Statistics**, [S.l.], n.15, p.161–175, 1988.
- MATLAB. **Matlab r2015a**. Matworks.
- MATTHIAS, A.; WOLFGANG, H.; HERBERT, W.; REINHARD, B.; CHRISTOPH, B. Adaptive AR Modeling of Nonstationary Time Series by Means of Kalman Filtering. **IEEE Transactions On Biomedical Engineering**, [S.l.], n.45, 1998.
- MCCULLOCH, W.; PITTS, W. A Logical Calculus of the Ideas Immanent in Nervous Activity. **Bulletin of Mathematical Biophysics**, [S.l.], n.5, p.115–133, 1943.
- MIELENZ, O.; KRUNER, A.; KOCHNER, H. Monitoramento avançado do nível de enchimento de ferro líquido e posição neutra dentro do forno. **MPT Edição Brasileira**, [S.l.], n.01, 2008.
- MORETTIN, A.; TOLOI, M. **Análise de Séries Temporais**. São Paulo, Brasil: Edgard Blucher, 2006.
- NIGHTINGALE, R.; TANZIL, F. Blast furnace hearth condition monitoring and taphole management techniques. **La Revue de Métallurgie, CIT**, [S.l.], n.06, p.533–540, 2001.
- NOCEDAL, J.; WRIGHT, S. J. **Numerical Optimization**. 2nd.ed. New York, USA: Springer, 2006.
- ONG, C. S.; HUANG, J.; TZENG, G. H. Model Identification of ARIMA Family Using Genetic Algorithms. **Applied Mathematics and Computation**, [S.l.], n.44, 2004.
- PEREIRA, A. **Projeto Ótimo de Pórticos Planos com Restrição à Flambagem**. Rio de Janeiro, Brasil: Pontifícia Universidade Católica do Rio de Janeiro, 2002.
- PETERS, M. **Determination of Liquid Level in the hearth of the Blast Furnace**. 2001 Ironmaking Conference Proceedings.
- PETERS, M. Continuous measurement of hot metal flow rates at the blast furnace. **Proc. 3rd International Conference on Science and Technology of Ironmaking**, [S.l.], p.477–481, 2002.
- POOS, A. Measurement techniques in blast furnace ironmaking and their benefits for industrial practice. **Proc. ECSC Workshop**, [S.l.], 1998.
- POST, J. **Hot Metal Flow in the Blast Furnace Hearth-Thermal and Carbon Dissolution Effects on Buoyancy, Flow and Refractory Wear**. Third International Conference on CFD in the Minerals and Process Industries CSIRO.
- PRICE, K. **An Introduction to Differential Evolution**. London: McGraw-Hill, 1999.

PRONIN, P. Variation in the electrical current voltage in the casting of blast furnace. **Izv. V.U.Z. Chernaya Metall**, [S.l.], n.05, 1985.

RADILOV, S. Electrical current in the blast furnace and its use for blast furnace monitoring. **STAL**, [S.l.], n.03, 1985.

REZENDE, S. O. **Sistemas Inteligentes: fundamentos e aplicações**. Barueri, Brasil: Editora Manole LTDA, 2003.

RUFF, W. Electricity in the blast furnace. **Stahl und Eisen**, [S.l.], n.01, 1927.

RUTHER, H.-P. Improvement of blast furnace process control by combined on-line monitoring of liquid level in the hearth and tapping flow rates. **Annual report, ECSC project 7210-PR-323**, [S.l.], 2003.

SHENOI, B. A. **Introduction to Digital Signal Processing and Filter Design**. 2.ed. Hoboken, USA: John Wiley and Sons, 2006.

SILVA, M. A. **Aplicação do Lagrangeano Aumentado em Otimização Estrutural com Restrições Dinâmicas**. São Paulo, Brasil: Escola Politécnica da Universidade de São Paulo, 1997.

SMIRNOV, N. Table for estimating the goodness of fit of empirical distributions. **Annals of Mathematical Statistics**, [S.l.], n.19, p.279–281, 1948.

SMITH, S. W. **The Scientist and Engineer's Guide to Digital Signal Processing**. 2.ed. San Diego, USA: California Technical Publishing, 1999.

SOUZA, R.; CAMARCO, M. **Análise e Previsão de Séries Temporais - Os Modelos ARIMA**. Rio de Janeiro, Brasil: Gráfica e Editora Regional, 2004.

STORN, R.; PRICE, K. Differential Evolution: a simple and efficient adaptive scheme for global optimization over continuous spaces. **International Computer Science Institute**, [S.l.], n.95, 1995.

STORN/CODE.HTML <http://www1.icsi.berkeley.edu/>.

<http://www1.icsi.berkeley.edu/storn/code.html>.

<http://www1.icsi.berkeley.edu/storn/code.html>.

TAN, L.; JIANG, J. **Digital Signal Processing - Fundamentals and Applications**. 2.ed. Oxford, United Kingdom: Academic Press - Elsevier, 2013.

THIEMANN, T. Optimisation of tapping strategy by advanced assessment of the liquid level in the hearth of the blast furnace. **Final report, ECSC project 7210-AA/142/321/820**, [S.l.], 2001.

THORNHILL, N.; HUANG, B.; ZHANG, H. Detection of multiple oscillations in control loops. **Journal of Process Control**, [S.l.], n.13, p.91–100, 2003.

TORRKULLA, J. Indicators of the internal state of the blast furnace hearth. **ISIJ International**, [S.l.], n.42, p.504–511, 2002.

TORRKULLA, J.; SAXEN, H. Model of the state of the blast furnace hearth. **ISIJ International**, [S.l.], n.40, p.438–447, 2000.

WEI, W. **Time Series Analysis**: univariate and multivariate methods. New York, USA: Person Education Inc, 2007.

WRIGHT, B.; ZULLI, P.; BIERBRAUER, F.; PANJKOVIC, V. **Assessment of Refractory Condition in a Blast Furnace Hearth Using Computational Fluid Dynamics**. Third International Conference on CFD in the Minerals and Process Industries CSIRO.

WWW.IMAGES.ENCARTA.MSN.COM. **www.images.encarta.msn.com**.
www.images.encarta.msn.com.

YANGMING, G.; LU, Z.; XIAOBIN, C.; CONGBAO, R.; ZHENGJUN, Z.; JIEZHONG, M. Time Series Adaptive Online Prediction Method Combined with Modified LS-SVR and AGO. **Mathematical Problems in Engineering**, [S.l.], 2012.

YAO, L.; SETHARES, W. A. Nonlinear Parameters Estimation via the Genetic Algorithm. **IEEE Transactions on Signal Processing**, [S.l.], n.42, p.927–935, 1994.

ZHANG, G. P.; PATUWO, B. E.; HU, M. Y. A simulation study of artificial neural networks for nonlinear time-series forecasting. **Computer & Operations Research**, [S.l.], n.28, p.381–396, 2001.

ZHANG, G. P.; QI, M. Neural network forecasting for seasonal and trend time series. **European Journal of Operational Research**, [S.l.], n.160, p.501–514, 2005.

APÊNDICE A REAÇÕES QUÍMICAS NO ALTO-FORNO

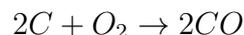
A principal reação para a operação de um alto-forno se dá entre o oxigênio do ar e o carbono do coque (FUJIHARA, 1994) que produz CO :



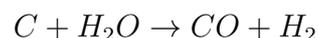
Se for considerada a região em frente à ventaneira, com excesso de carbono e elevada temperatura, tem-se:



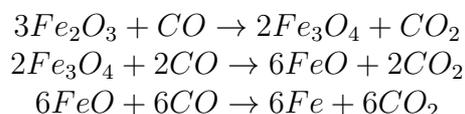
Pela lei de Hess (LEICESTER, 1951), tem-se:



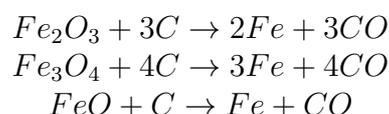
O produto de reação (CO) irá ascender pelo alto-forno, reduzindo os óxidos. O calor gerado por essa reação irá aquecer e fundir os sólidos. Em frente à ventaneira, devido à presença de umidade no ar de sopro, tem-se a seguinte reação gerando um gás redutor diferente, que é o H_2 :



Além dessas equações, tem-se em frente às ventaneiras, a decomposição dos combustíveis auxiliares: óleo, alcatrão e/ou carvão gerando CO , H_2 e N_2 em porcentagens variáveis de acordo com as condições operacionais. O minério-de-ferro, de modo geral, se encontra na forma de óxido. Assim, as reações de redução podem ser descritas da seguinte maneira:



Estas reações são chamadas de reações de redução indireta e essas reações ocorrem inteiramente no estado sólido, sendo necessárias altas temperaturas devido a certas reações serem endotérmicas. A redução do óxido-de-ferro pode se dar também através do redutor carbono, conforme abaixo:



Estas reações são chamadas de reações de redução direta. Também são reações endotérmicas e evoluem com o aumento de temperatura. Devido à condição de aquecimento da carga e redução pelo CO , basicamente não há ocorrência da redução direta dos óxidos

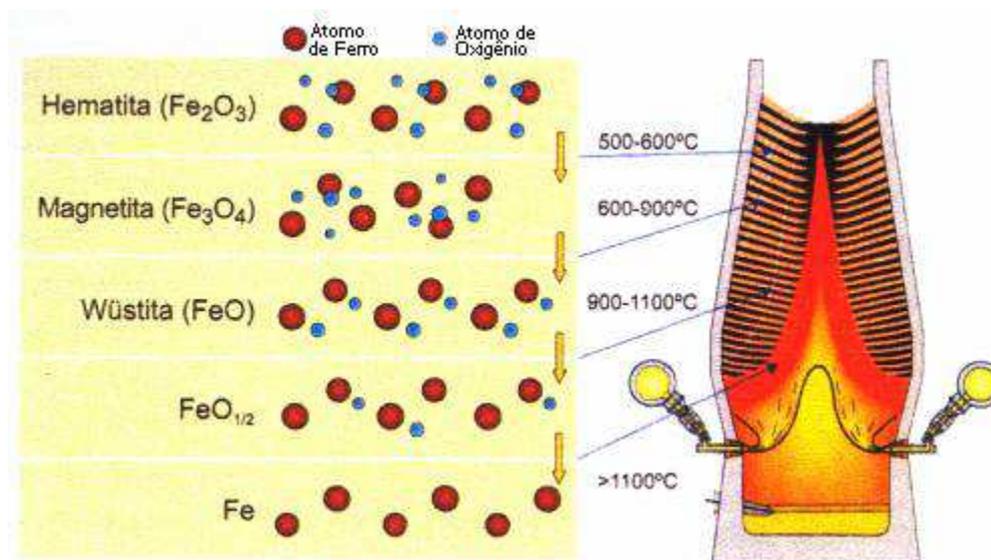


Figura 66: Representação esquemática dos óxidos de ferro em função da temperatura no interior do alto-forno.

Fe_2O_3 e Fe_3O_4 . Já a redução do FeO , devido as temperaturas envolvidas, se dá na região mais inferior da cuba, conforme pode ser visto na Figura 66.

As composições do gusa e da escória que saem são apresentadas na Tabela 15. Na aciaria, o gusa é ajustado para que se enquadre nos requerimentos metalúrgicos. O processo de refino consiste, basicamente, de dois passos: dessulfuração (retirada de enxofre do gusa) pela inserção de carbetto, cal ou magnésio e a remoção de carbono, silício, manganês e fósforo pelo sopro de oxigênio.

Tabela 15: Composição típica do gusa e da escória

	Gusa		Escória		
	Típico(%)	Faixa(%)		Típico(%)	Faixa(%)
Fe	94.5	93 – 96	CaO	40.0	34 – 42
C	4.5	3.8 – 4.7	MgO	7.0	5 – 12
Si	0.4	0.3 – 0.47	SiO_2	36.0	28 – 36
Mn	0.3	0.2 – 0.5	Al_2O_3	12.0	8 – 20
P	0.07	0.03 – 1	Soma	96.5	96 – 97
Ti	0.5	0.3 – 0.7	S	1.0	0.8 – 1.2

Fonte:(GEERDES et al., 2007)

A composição final do gusa é o resultado de um processo complexo de interação metal-escória. Os vários elementos se dividem entre a escória e gusa líquidos. A dispersão de um elemento pelas duas fases depende da composição da escória e da temperatura do processo. Os óxidos de silício, manganês e fósforo são reduzidos através da reação de redução direta. A quantidade de silício e de manganês no gusa são indicativos do estado térmico do forno podendo ser utilizados para análise da consistência do processo. A escória é formada a partir da carga, da cinza do coque e dos combustíveis auxiliares. Quatro componentes principais compõem aproximadamente 96% da escória (SiO_2 , MgO , CaO e Al_2O_3). O balanço químico (para formar 100%) vem de componentes tais como o óxido de manganês (MnO), enxofre (S), titânia (TiO_2), potassa (K_2O) e óxido

de fósforo (P_2O_5).

APÊNDICE B DEFINIÇÕES GERAIS EM ESTATÍSTICA

B.1 Esperança de Processos Estocásticos

Suponha uma amostra observada de tamanho t da variável aleatória Z_t : Z_1, Z_2, \dots, Z_t . Considere agora uma coleção de t variáveis independentes e identicamente distribuídas $a_T, a_1, a_2, \dots, a_t$ com, $a_t \sim N(0, \sigma^2)$. Sendo N a distribuição normal e σ o desvio padrão.

Isto se refere a uma amostra de tamanho T de um processo ruído gaussiano branco. A amostra observada acima representa T números particulares, mas o conjunto de T números representados é somente uma possível realização do processo estocástico que gerou a amostra. Se for imaginada a observação do processo por um período infinito de tempo, tem-se a seguinte seqüência:

$$\{Z_t\}_{t=-\infty}^{\infty} = \{\dots, Z_{-1}, Z_0, Z_1, Z_2, \dots, Z_T, Z_{T+1}, Z_{T+2}, \dots\} \quad (\text{B.1})$$

Esta seqüência infinita $\{Z_t\}_{t=-\infty}^{\infty}$ pode ser vista como uma simples realização de um processo de séries temporais. Por exemplo, se colocarmos um computador para gerar uma seqüência infinita com $N(0, \sigma^2), \{a_t^{(1)}\}_{t=-\infty}^{\infty}$ e um segundo computador gerando uma nova seqüência $\{a_t^{(2)}\}_{t=-\infty}^{\infty}$ veremos que estas duas serão realizações independentes de um processo tipo ruído branco, que será explicado posteriormente.

Imagine uma seqüência de tamanho I gerada a partir das seqüências $\{Z_t^{(1)}\}_{t=-\infty}^{\infty}, \{Z_t^{(2)}\}_{t=-\infty}^{\infty}, \dots, \{Z_t^{(I)}\}_{t=-\infty}^{\infty}$, e considere a seleção de observações associadas ao tempo t de cada seqüência: $\{Z_t^{\{1\}}, Z_t^{\{2\}}, \dots, Z_t^{\{I\}}\}$. Esta seqüência pode ser descrita como uma amostra de I realizações de uma variável Z_t . Esta variável aleatória possui uma determinada densidade de probabilidade, denotada como $f_{Z_t}(Z_t)$, que é conhecida como densidade condicional de Z_t . Por exemplo, para um processo gaussiano ruído branco, esta densidade é dada por:

$$f_{Z_t}(Z_t) = \frac{1}{\sqrt{2\pi\sigma}} \exp\left\{-\frac{Z_t^2}{2\sigma^2}\right\} \quad (\text{B.2})$$

A esperança da t -ésima observação de uma série temporal refere-se à média da sua distribuição de probabilidade, ou seja:

$$E(Z_t) = \int_{-\infty}^{\infty} Z_t f_{Z_t}(Z_t) dZ_t \quad (\text{B.3})$$

Pode-se referir à esperança $E(Z_t)$ como a média incondicional de Z_t ou μ_t :

$$E(Z_t) = \mu_t \quad (\text{B.4})$$

A variância da variável aleatória Z_t (denotada por γ_{0t}) é similarmente definida como (HAMILTON, 1994):

$$\gamma_{0t} = E(Z_t - \mu_t)^2 = \int_{-\infty}^{\infty} (Z_t - \mu_t)^2 f_{Z_t}(Z_t) dZ_t \quad (\text{B.5})$$

correspondendo à dispersão das amostras em relação à seu valor esperado μ_t .

B.2 Autocovariância

A autocovariância de lag k é definida como a covariância entre Z_t e Z_{t+k} , ou seja:

$$\gamma_k = Cov(Z_t, Z_{t+k}) = E[(Z_t - \mu_t)(Z_{t+k} - \mu_{t+k})] \text{ para } k=0,1,2,\dots \quad (\text{B.6})$$

Note que, $\gamma_0 = Var(Z_t) = \sigma^2$ pois $Cov(Z_t, Z_t) = Var(Z_t)$. A equação (??) pode ser descrita como a covariância de Z em relação a seus valores anteriores, daí vem o termo autocovariância (HAMILTON, 1994).

B.3 Autocorrelação

A autocorrelação (*Auto-Correlation Function* - ACF) é a autocovariância padronizada pela variância e é utilizada para medição do comprimento e a memória de um processo, ou seja, a extensão para qual o valor tomado no instante t depende de um valor tomado no instante $t - k$. A autocorrelação de lag k é dada por:

$$\rho_k = \frac{Cov(Z_t, Z_{t+k})}{Var(Z_t)} = \frac{\gamma_k}{\sigma^2} \text{ para } k=0,1,2,\dots \quad (\text{B.7})$$

B.4 Autocorrelação Parcial

A autocorrelação parcial (*Partial Auto-Correlation Function* - PACF) é a autocorrelação medida entre duas observações seriais Z_t e Z_{t+k} eliminando-se a dependência dos termos intermediários, $Z_{t+1}, Z_{t+2}, \dots, Z_{t+k-1}$. A autocorrelação parcial de defasagem k é dada por:

$$\rho_{kk} = (Z_t, Z_{t+k} \mid Z_{t+1}, \dots, Z_{t+k-1}). \quad (\text{B.8})$$

B.5 Estacionaridade

Um processo é chamado fracamente estacionário ou estacionário de segunda ordem se a média μ_t e a autocovariância γ_k satisfazem: $E(Z_t) = \mu_t = \mu$ e $E[(Z_t - \mu_t)(Z_{t+k} - \mu_{t+k})]$ para todo t .

Um processo é chamado estritamente estacionário se, para quaisquer valores de k_1, k_2, \dots, k_n , a soma das distribuições de $(Z_t, Z_{t+k_1}, Z_{t+k_2}, \dots, Z_{t+k_n})$ depende somente dos intervalos que separam os tempos (k_1, k_2, \dots, k_n) e não do tempo propriamente dito (t). Isto significa que todas as distribuições unidimensionais são invariantes sob translações do tempo, logo: $E(Z_t) = \mu$ e $Var(Z_t) = \mu^2$ para todo t .

Como a grande maioria dos processos são fracamente estacionários, a partir de agora chamaremos estes processos de simplesmente processos estacionários. Alguns processos estacionários podem ser diferenciados para a obtenção de processos estritamente estacionários (HAMILTON, 1994).

B.6 Normalidade

Uma distribuição normal ou Gaussiana é:

$$f(x, \mu, \sigma) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \quad (\text{B.9})$$

O parâmetro μ na equação acima é a média ou esperança da distribuição. O parâmetro σ é o desvio-padrão; A variância σ^2 .

Se $\mu = 0$ e $\sigma = 1$, a distribuição é chamada distribuição normal padrão ou distribuição normal unitária.

B.7 Ruído Branco

Dizemos que um sinal $\{a_t, t \in Z\}$ é ruído branco discreto se as variáveis aleatórias a_t são não-correlacionadas, isto é, $Cov\{a_t, a_s\} = 0$ para $t \neq s$. Uma seqüência de variáveis aleatórias independentes e não-correlacionadas que possui as seguintes características: $E(a) = 0$ e $E(a^2) = \sigma^2$ onde $a_T \sim N(0, \sigma^2)$ é chamado de ruído branco gaussiano (MORETTIN; TOLOI, 2006).

B.8 Densidade Espectral de Potência

A transformada de Fourier é tipicamente utilizada para decompor um sinal nas suas componentes em frequência e suas amplitudes. Entretanto esta ferramenta matemática é fortemente influenciada pelo ruído existente no sinal no tempo ocasionando falsos picos de potência no espectro, o que dificulta substancialmente sua análise.

A *Power Spectral Density* - PSD é calculada a partir da transformada de Fourier rápida (FFT) aplicada sobre o sinal autocorrelacionado, o que minimiza o efeito do ruído no espectro. A PSD pode ser descrita matematicamente por,

$$\Phi_x(f) = \sum_{k=1}^N r_{xx}(k) e^{-j2\pi f k} \quad (\text{B.10})$$

onde f é a frequência, e N é o número de atrasos de $r_{xx}(k)$. A autocorrelação $r_{xx}(k)$ é descrita matematicamente por,

$$r_{xx}(k) = \frac{\sum_{t=1}^{N-k} [x(t) - \bar{x}][x(t+k) - \bar{x}]}{\sum_{t=1}^N [x(t) - \bar{x}]^2} \quad (\text{B.11})$$

onde \bar{x} é o valor médio para as N amostras, e k é o atraso.